

19 JUN 1998

09/09-538

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: St George's Hospital Medical School
- (B) STREET: Cranmer Terrace
- (C) CITY: London
- (E) COUNTRY: United Kingdom
- (F) POSTAL CODE (ZIP): SW17 0RE

(ii) TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES IN
PATHOGENIC MYCOBACTERIA AND THEIR USE AS DIAGNOSTICS,
VACCINES AND TARGETS FOR CHEMOTHERAPY

(iii) NUMBER OF SEQUENCES: 41

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(v) CURRENT APPLICATION DATA:

APPLICATION NUMBER: WO PCT/GB96/03221

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GATCCA	ACTA	AACCC	GATGG	AACCC	CGCGC	AACTA	TATGG	ACGTCT	CCGC	GCTAC	GCAGT	60
TGGGT	TGGCG	CCGCG	GAATC	GCACT	GAAAG	AGGGC	ATCGA	TGCAAC	GGTG	TCGTG	GTACC	120
GCACAA	AATGC	CGATG	CCGTG	AGGAG	GTAAA	GCTGC	GGGCC	GGCCG	ATGTT	ATCCCT	CCGG	180
CCGGAC	GGGT	AGGGC	GACCT	GCCAT	CGAGT	GGTAC	GGCAG	TCGCCT	GGCC	GGCGAG	GGCGC	240
ATGGC	CTATG	TGAGT	ATCCC	ATAGC	CTGGC	TTGGC	TCGCC	CCTAC	GCATT	ATCAGT	TGAC	300

CGCTTTCGCG CCACGTCGCA GGCTTGCGGC AGCATCCCGT TCAGGTCTCC TCATGGTCCG 360
 GTGTGGCACG ACCACGCAAG CTCGAACCGA CTCGTTTCCC AATTTGCGAT GCTAATATCG 420
 CTCGATGGAT TTTTTCGCA ACGCCGGCTT GATGGCTCGT AACGTTAGCA CCGAGATGCT 480
 GCGCCACTCC GAACGAAAGC GCCTATTAGT AAACCAAGTC GAAGCATACG GAGTCAACGT 540
 TGTTATTGAT GTCGGTGCTA ACTCCGGCCA GTTCGGTAGC GCTTTGCGTC GTGCAGGATT 600
 CAAGAGCCGT ATCGTTTCCT TTGAACCTCT TTCGGGGCCA TTTGCGCAAC TAACGCGCAA 660
 GTCGGCATCG GATC 674

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

GATCCGATGC CGACTTGCGC GTTAGTTGCG CAAATGGCCC CGAAAGAGGT TCAAAGGAAA 60
 CGATACGGCT CTTGAATCCT GCACGACGCA AAGCGCTACC GAACTGGCCG GAGTTAGCAC 120
 CGACATCAAT AACAACGTTG ACTCCGTATG CTTCGACTTG GTTTACTAAT AGGCGCTTTC 180
 GTTCGGAGTG GCGCAGCATC TCGGTGCTAA CGTTACGAGC CATCAAGCCG GCGTTGCGCA 240
 AAAAATCCAT CGAGCGATAT TAGCATGCGA AATTGGGAAA CGAGTCGGTT CGAGCTTGCG 300
 TGGTCGTGCC ACACCGGACC ATGAGGAGAC CTGAACGGGA TGCTGCCGCA AGCCTGCGAC 360
 GTGGCGCGAA AGCGGTCAAC TGATAATGCG TAGGGGCGAG CCAAGCCAGG CTATGGGATA 420
 CTCACATAGG CCATGCGCCT CGCCGGCCAG GCGACTGCCG TACCACTCGA TGGCAGGTCG 480
 CCCTACCCGT CCGGCCGGAG GGATAACATC GGCCGGCCCG CAGCTTTACC TCCTCACGGC 540
 ATCGGCATTT GTGCGGTACC ACGACACCGT TGCATCGATG CCCTCTTTCA GTGCGATTCTG 600
 CGGGCGCCAA CCCAACTGCG TAGCGCGGAG ACGTCCAATA GTTTGCGCGG GGTTCATCG 660

GGTTTAGTTG GATC

674

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7995 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

GAATTCTGGG TTGGAGACGA CGTCGAACTC CTGGTCGGTC TTGCTTCGAA TGATCGCTGT	60
GATCTGGTCG GCGGTGCCGA CAGGAACCGT CCACTTGTCTG ACGATCACCT TGTACCGGTC	120
GATGTATGAC CCAATGTCGT CCGCAACCGA GAAGACGTAC GTCAGGTCCG CCGCCCCGCT	180
TTCACCCATG GGCCTCGGGA CGGCGATGAA AATGACGTCC GCGTGCTCGA TTCCGCGTTG	240
CCGGTCGGTG GTGAAGTCAA TCAGCCCGTT CTCACGGTTC CTCGCAATCA ACTCCCAACC	300
CGGGCTCGAA AATCGGGACA CTGCCTGCGA GGAGCAAATC GATCTTGGCC TGATCGATAT	360
CGACACAGAC GACATCGTTG CCGCTATCCG CGAGACAGGC GCCCGTGACG AGGCCTACAT	420
AGCCTGATCC GACCACCGAA ATTTTCAAGA TGACCCCTTC AAGTCCCCGA TCGGTCGACG	480
ACCATACTGC CGCAACTCTG TACCCTCCGT GGGTAATTCG CATGTCGCGT TCGTAAGGAG	540
CAGCCAGCGA GTCGGGGACG TTCGGTGAGA GAGTCGCAGG ACTACGAGGT TGCCGGTGCG	600
ATACATCACA GTGTTGCGTC TGTCGGCAAC GATGCAGCAA GAACCCACGG GGCAGCCCTG	660
AACTGCGCGC ATGACCGGTC CTTGTCCTGG CACCTTTGAT CGGCCACCGC TTCCATGCGA	720
ACATGACCGG AATCCATAGC GCGTGGTCAA GCAGCGGGGA GGTAGACGTC GGTGTCATCT	780
GCTCCAACCG TGTCGGTGAT AACGATTTCTG CTGAACGATC TCGAGGGATT GAAAAGCACC	840
GTGGAGAGCG TTCGCGCGCA GCGCTATGGG GGGCGAATCG AGCACATCGT CATCGACGGT	900
GGATCGGGCG ACGCCGTCGT GGAGTATCTG TCCGGCGATC CTGGCTTTGC ATATTGGCAA	960
TCTCAGCCCCG ACAACGGGAG ATATGACGCG ATGAATCAGG GCATTGCCCA TTCGTCGGGC	1020

CCTCGATGCG ATCCGCGACT GGGGCTACGC GCCCGAATAT GTCGAGGGGA TGTGGAGGAT	2520
GTTGCAAGCG CCTGAACCTG ATGACTACGT CCTGGCGACA GGGCGTGGTT ACACCGTACG	2580
TGAGTTCGCT CAAGCTGCTT TTGACCATGT CGGGCTCGAC TGGCAAAAGC GCGTCAAGTT	2640
TGACGACCGC TATTTGCGTC CCACCGAGGT CGATTGCTA GTAGGAGATG CCGACAAGGC	2700
GGCCCAGTCA CTCGGCTGGA AAGCTTCGGT TCATACTGGT GAACTCGCGC GCATCATGGT	2760
GGACGCGGAC ATCGCCGCGT TGGAGTGCGA TGGCACACCA TGGATCGACA CGCCGATGTT	2820
GCCTGGTTGG GGCAGAGTAA GTTGACGACT ACACCTGGGC CTCTGGACCG CGCAACGCCC	2880
GTGTATATCG CCGGTCATCG GGGGCTGGTC GGCTCAGCGC TCGTACGTAG ATTTGAGGCC	2940
GAGGGGTTCA CCAATCTCAT TGTGCGATCA CGCGATGAGA TTGATCTGAC GGACCGAGCC	3000
GCAACGTTTG ATTTTGTGTC TGAGACAAGA CCACAGGTGA TCATCGATGC GGCCGCACGG	3060
GTCGGCGGCA TCATGGCGAA TAACACCTAT CCCGCGGACT TCTTGTCCGA AAACCTCCGA	3120
ATCCAGACCA ATTTGCTCGA CGCAGCTGTC GCCGTGCGTG TGCCGCGGCT CCTTTTCCTC	3180
GGTTCGTCAT GCATCTACCC GAAGTACGCT CCGCAACCTA TCCACGAGAG TGCTTTATTG	3240
ACTGGCCCTT TGGAGCCCAC CAACGACGCG TATGCGATCG CCAAGATCGC CGGTATCCTG	3300
CAAGTTCAGG CGGTTAGGCG CCAATATGGG CTGGCGTGGA TCTCTGCGAT GCCGACTAAC	3360
CTCTACGGAC CCGGCGACAA CTTCTCCCCG TCCGGGTCGC ATCTCTTGCC GGCCTCATC	3420
CGTCGATATG AGGAAGCCAA AGCTGGTGGT GCAGAAGAGG TGACGAATTG GGGGACCGGT	3480
ACTCCGCGGC GCGAACTTCT GCATGTCGAC GATCTGGCGA GCGCATGCCT GTTCCTTTTG	3540
GAACATTTTCG ATGGTCCGAA CCACGTCAAC GTGGGCACCG GCGTCGATCA CAGCATTAGC	3600
GAGATCGCAG ACATGGTCGC TACAGCGGTG GGCTACATCG GCGAAACACG TTGGGATCCA	3660
ACTAAACCCG ATGGAACCCC GCGCAAACCTA TTGGACGTCT CCGCGCTACG CGAGTTGGGT	3720
TGGCGCCCGC GAATCGCACT GAAAGACGGC ATCGATGCAA CGGTGTCGTG GTACCGCACA	3780
AATGCCGATG CCGTGAGGAG GTAAAGCTGC GGGTCGGCCG ATGTTATCCC TCCGGCCGGA	3840
CGGGTGGGGC GACCTGCCGT CGAGTGGTAC GGCAGTCGCC TGGCCGGCGA GGCCTGTGGC	3900

ATGCCGGACC	TTTCGACCTC	GACCGCCTCC	TATTTGAGAC	GAATTTGTGC	CACCAATCGA	5400
TCTTTTACCG	CCGTGAGCTT	TTCGACGGCA	TCGGCCCTTA	CAACCTGCGC	TACCGAGTCT	5460
GGGCGGACTG	GGACTIONAAT	ATTCGCTGCT	TCTCCAACCC	GGCGCTGATT	ACCCGCTACA	5520
TGGACGTCGT	GATTTCCGAA	TACAACGACA	TGACCGGCTT	CAGCATGAGG	CAGGGGACTG	5580
ATAAAGAGTT	CAGAAAACGG	CTGCCAATGT	ACTTCTGGGT	TGCAGGGTGG	GAGACTTGCA	5640
GGCGCATGCT	GGCGTTTTTG	AAAGACAAGG	AGAATCGCCG	TCTGGCCTTG	CGTACGCGGT	5700
TGATAAGGGT	TAAGGCCGTC	TCCAAAGAAC	GAAGCGCAGA	ACCGTAGTCG	CGGATCCACA	5760
TTGGACTTCT	TTAACGCGTT	TGCGTCCTGA	TCCACCTTTC	AAGCCCGTTC	CGCGTAACGC	5820
GGCGCGCAGA	GAGTGGTCGC	ATATCGCATC	ACTGTTCTCG	TGCCAGTGCT	TGGAAAGCGT	5880
CGAGCACTCT	GGTTCGCGTT	CTTGACGTTT	GCGCCCGCTC	CTAGAGGTAG	CGTGTCACGT	5940
GACTGAAGCC	AATGAGTGCA	ACTCGGCGTC	GCGAAAGGTT	TCAGTCGCGG	TTGAGCAAGA	6000
CACCGCAAGA	CTACTGGAGT	GCGTGACAAA	GCGCCTCCAG	CTCGCGGCTG	AAAGCGGATG	6060
CAAAGGGATT	CGAAGCTTGA	GCAACATGCG	AAGGGGAGAA	CGGCCTATGA	GGCTGGGACA	6120
GGTTTTCGAT	CCGCGCGCGA	ATGCACTGTC	AATGGCCAAG	TAGAAGTCCC	CGCTGGTGGC	6180
CAGCAGAAGT	CCCCACTCCG	CTGCGGGTGG	TTGGCTAATT	CTTGGCGGCT	CCCTTCTTGT	6240
GGTCGGCGTG	GCGCATCCGG	TAGGACTCGC	CGGAGGTGAC	GACGATGCTG	GCGTGGTGCA	6300
GCAGCCGATC	GAGGATGCTG	GCGGCGGTGG	TGTGCTCGGG	CAGGAATCGC	CCCCATTGTT	6360
CGAAGGGCCA	ATGCGAGGCG	ATGGCCAGGG	AGCGGCGCTC	GTAGCCGGCA	GCCACGAGCC	6420
GGAACAACAG	TTGAGTCCCG	GTGTCGTCGA	GCGGGGCGAA	GCCGATCTCG	TCCAAGATGA	6480
CCAGATCCGC	GCGGAGCAGG	GTGTCGATGA	TCTTGCCGAC	GGTGTGTGTC	GCCAGGCCGC	6540
GGTAGAGGAC	CTCGATCAGG	TCGGCGGCGG	TGAAGTAGCG	GACTTTGAAT	CCGGCGTGGA	6600
CGGCAGCGTG	CCCGCAGCCG	ATGAGCAGGT	GACTTTTGCC	CGTACCAGGT	GGGCCAATGA	6660
CCGCCAGGTT	CTGTTGTGCC	CGAATCCATT	CCAGGCTCGA	CAGGTAGTCG	AACGTGGCTG	6720
CGGTGATCGA	CGATCCGGTG	ACGTCGAACC	CGTCGAGGGT	CTTGGTGACC	GGGAAGGCTG	6780

CGGCCTTGAG ACGGTTGGCG GTGTTGGAGG CATCGCGGGC AGCGATCTCG GCCTCAACCA 6840
 ACGTCCGCAG GATCTCCTCC GGTGTCCAGC GTTGCCTCTT GGCCTACTGC AACACCTCGG 6900
 CGGCGTTGCG GCGCACCGTG GCCAGCTTCA ACCGCCGCAG CGCCGCGTCA AGGTCAGCAG 6960
 CCAGCGGTGC CGCCGAGGAC GGTGCCACCG GCTTGGCAGC GGTGGTCATG AGGCCGTCCC 7020
 GTCGGTGGTG TTGATCTTGT AGGCCTCCAA CGAGCGGGTC TCGACGGTGG GCAGATCGAG 7080
 CACGAGTGCG TCGCCGGCGG GGCGGGGTTG TGGGGTGCCG GCGCCGGCGG CCAGGATCGA 7140
 GCGCACGTCG GCAGCGCGGA ACCGGCGAAA CGCAACCGCC CGGCGCAGCG CGTCAATCAA 7200
 AGCCTGTTTC CCGTGGGCGG CGCCAAGGCC GAGCAGAATG TCGAGTTCGG ATTCAGTCG 7260
 GGTGTTGCCG ATCGCAGCAG CACCGACGAG GAACTGCTGC GCTTCGGTTC CCAATGCGCA 7320
 GAATCGTTTC TCTGCTTGGG TTTTCGGGCG AGGACCACGC GAGGGTGCGG GTCTGGGTCC 7380
 GTCGTAGTGT TCATCGAGGA TGGACACCTC ACCTGGGCTG ACGAGCTCGT GCTCGGCCAC 7440
 GATCACACCG GTCGCAGGTT CCAACAGGAT CAGGGCGCCA TGATCGACCA CCACCGCCAC 7500
 GGTGGCACCG ACGAGCCGCT GAGGCACCGA GTAACGAGCT GAGCCGTAAC GGATGCACGA 7560
 GAGGCCGTCG ACCTTACGGC GCACCGACCC CGAGCCGATC GTCGGCCGCA GCGAGGGCAG 7620
 CTCCTCAAG ACGGTGCGCT CGTCAACCAA GCGATCGTTG GGCACGGCGC AGATCTCCGA 7680
 GTGGACCGTG GCATTGACCT CGGCGACCA TAGTTGCGCC TGGGCGTTGA GGGCACGTAG 7740
 GTCGACCTGC TCACCGGCTA ACGCAGCTTC GGTCAGCAGC GGCACCGCAA GGTCGTCCTG 7800
 AGCGTAGCCA CAGAGGTTCT CCACGATGCC CTTGATTGC GGATCCGCAC CGTGGCAGAA 7860
 GTCCGGAACG AAGCCATAGT GGGACGCGAA TCGCACATAA TCCGGTGTTG GAACAACAAC 7920
 ATTGGCGACG ACACCACCTT TGAGGCAGCC CATCCGGTCG GCCAGGATCT TGGCCGGAAC 7980
 CCCACCGATC GCCTC 7995

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 4435 base pairs

(B) TYPE: nucleic acid

GAGTCGGCAT CGGATCCACT ATGGGAGTGT CACCAGTATG CCCTAGGCGA CGCCGATGAG	2760
ACGATTACCA TCAATGTGGC AGGCAATGCG GGGGCAAGTA GTTCCGTGCT GCCGATGCTT	2820
AAAAGTCATC AAGATGCCTT TCCTCCCGCG AATTATATTG GCACCGAAGA CGTTGCAATA	2880
CACCGCCTTG ATTCGGTTGC ATCAGAATTT CTGAACCCTA CCGATGTTAC TTTCTGAAG	2940
ATCGACGTAC AGGGTTTCGA GAAGCAGGTT ATCGCGGGCA GTAAGTCAAC GCTTAACGAA	3000
AGCTGCGTCG GCATGCAACT CGAACTTTCT TTTATTCCGT TGTACGAAGG TGACATGCTG	3060
ATTCATGAAG CGCTTGAAC TGTCTATTCC CTAGGTTTCA GACTGACGGG TTTGTTGCCC	3120
GGATTTACGG ATCCGCGCAA TGGTCGAATG CTTCAAGCTG ACGGCATTTT CTTCCGTGGG	3180
GACGATTGAC ATAAATGCTT GCGTCGGCAC CCTGCCGTA TCCAAACGGG CGATCTGGTG	3240
AGCCGGCCTC CCGGGCACCT AATCGACTAT CTAAATTGAG GCGGCCGCGA CGTGCGGCAC	3300
GAACAGGTGG CCGGCTGCTA GCGTTACACA CGTCATGACT GCGCCAGTGT TCTCGATAAT	3360
TATCCCTACC TTCAATGCAG CGGTGACGCT GCAAGCCTGC CTCGGAAGCA TCGTCGGGCA	3420
GACCTACCGG GAAGTGGAAG TGGTCCTTGT CGACGGCGGT TCGACCGATC GGACCCTCGA	3480
CATCGCGAAC AGTTTCCGCC CGGAACTCGG CTCGCGACTG GTCGTTTACA GCGGGCCCGA	3540
TGATGGCCCC TACGACGCCA TGAACCGCGG CGTCGGCGTA GCCACAGGCG AATGGGTACT	3600
TTTTTTAGGC GCCGACGACA CCCTCTACGA ACCAACCACG TTGGCCCAGG TAGCCGCTTT	3660
TCTCGGCGAC CATGCGGCAA GCCATCTTGT CTATGGCGAT GTTGTGATGC GTTCGACGAA	3720
AAGCCGGCAT GCCGGACCTT TCGACCTCGA CCGCCTCCTA TTTGAGACGA ATTTGTGCCA	3780
CCAATCGATC TTTTACCGCC GTGAGCTTTT CGACGGCATC GGCCCTTACA ACCTGCGCTA	3840
CCGAGTCTGG GCGGACTGGG ACTTCAATAT TCGCTGCTT TCCAACCCGG CGCTGATTAC	3900
CCGCTACATG GACGTCGTGA TTTCCGAATA CAACGACATG ACCGGCTTCA GCATGAGGCA	3960
GGGGACTGAT AAAGAGTTCA GAAAACGGCT GCCAATGTAC TTCTGGGTTG CAGGGTGGGA	4020
GACTTGACAGG CGCATGCTGG CGTTTTTGAA AGACAAGGAG AATCGCCGTC TGGCCTTGCG	4080
TACGCGGTTG ATAAGGGTTA AGGCCGTCTC CAAAGAACGA AGCGCAGAAC CGTAGTCGCG	4140

(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

TTCTACTGCC TGACCTGAGC AGCGCCGAGG CGCGCAGCGC GATCACTGCG ACCTGAATGG	60
CCAGGTGGAA AGCGCCACCG ATCCCGGCAC CGAGTGCCTG ACGATTCGGA TCCCTTGCAC	120
CACAACGAGA GTGAGACCGC CATGATGACG AAATATCGGC TGGGCGGAGT CAACGCCGGA	180
GTGACAAAAG TGAGAACCCG GTGAAGCGAG CGCTTATAAC AGGGATCACG GGGCAGGATG	240
GTTCTACCT CGCCGAGCTA CTA CTGAGCA AGGGATACGA GGTTCACGGG CTCGTTCGTC	300
GAGCTTCGAC GTTTAACACG TCGCGGATCG ATCACCTCTA CGTTGACCCA CACCAACCGG	360
GCGCGCGCTT GTTCTTGCAC TATGCAGACC TCACTGACGG CACCCGGTTG GTGACCTGC	420
TCAGCAGTAT CGACCCGGAT GAGGTCTACA ACCTCGCAGC GCAGTCCCAT GTGCGCGTCA	480
GCTTTGACGA GCCAGTGCAT ACCGGAGACA CCACCGGCAT GGGATCGATC CGACTTCTGG	540
AAGCAGTCCG CCTTTCTCGG GTGGACTGCC GGTTCATCA GGCTTCCTCG TCGGAGATGT	600
TCGGCGCATC TCCGCCACCG CAGAACGAAT CGACGCCGTT CTATCCCCGT TCGCCATACG	660
GCGCGGCCAA GGTCTTCTCG TACTGGACGA CTCGCAACTA TCGAGAGGCG TACGGATTAT	720
TCGCAGTGAA TGGCATCTTG TTCAACCATG AGTCCCCCGG GCGCGGCGAG ACTTTCGTGA	780
CCCGAAAGAT CACGCGTGCC GTGGCGCGCA TCCGAGCTGG CTGCCAATCG GAGGTCTATA	840
TGGGCAACCT CGATGCGATC CGCGACTGGG GCTACGCGCC CGAATATGTC GAGGGGATGT	900
GGAGGATGTT GCAAGCGCCT GAACCTGATG ACTACGTCCT GGCACAGGG CGTGATTACA	960
CCGTACGTGA GTTCGCTCAA GCTGCTTTTG ACCACGTCGG GCTCGACTGG CAAAAGCACG	1020
TCAAGTTTGA CGACCGCTAT TTGCGCCCCA CCGAGGTCGA TTCGCTAGTA GGAGATGCCG	1080
ACAGGGCGGC CCAGTCACTC GGCTGGAAAG CTTCGGTTCA TACTGGTGAA CTCGCGCGCA	1140
TCATGGTGGA CGCGGACATC GCCGCGTCGG AGTGCGATGG CACACCATGG ATCGACACGC	1200
CGATGTTGCC TGGTTGGGGC GGAGTAAGTT GACGACTACA CCTGGGCCTC TGGACCGCGC	1260

GATCCACATT GGACTTCTTT AACGCGTTTG CGTCCTGATC CACCTTTCAA CCCC GTTCCG 4200
 CGTGACGCGG CGCGCAGAGA GTGGTCGCAT ATCGCGTCAC TGTTCTCGTG CCAGTGCTTG 4260
 GAAAGCGTCG AGCACTCTGG TTCGCGTTCT TGACGTTTCG GCGCGCCCT AGAGGTAGCG 4320
 TGTCACGTGA CTGAAGCCAA TGAGTGCAAC TCGGCGTCGC GAAAGGTTTC AGTCGCGGTT 4380
 GAGCAAGACA CCGCAAGACT ACTGGAGTGC GTGCACAAGC GCCTCCAGCT CACGG 4435

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 378 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..375

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATG ATC GCT GTG ATC TGG TCG GCG GTG CCG ACA GGA ACC GTC GAC TTG	48
Met Ile Ala Val Ile Trp Ser Ala Val Pro Thr Gly Thr Val Asp Leu	
1 5 10 15	
TCG ACG ATC ACC TTG TAC CGG TCG ATG TAT GAC CCA ATG TCG TCC GCA	96
Ser Thr Ile Thr Leu Tyr Arg Ser Met Tyr Asp Pro Met Ser Ser Ala	
20 25 30	
ACC GAG AAG ACG TAC GTC AGG TCC GCC GCC CCG CTT TCA CCC ATG GGC	144
Thr Glu Lys Thr Tyr Val Arg Ser Ala Ala Pro Leu Ser Pro Met Gly	
35 40 45	
GTC GGG ACG GCG ATG AAA ATG ACG TCC GCG TGC TCG ATT CCG CGT TGC	192
Val Gly Thr Ala Met Lys Met Thr Ser Ala Cys Ser Ile Pro Arg Cys	
50 55 60	
CGG TCG GTG GTG AAG TCA ATC AGC CCG TTC TCA CGG TTC CTC GCA ATC	240
Arg Ser Val Val Lys Ser Ile Ser Pro Phe Ser Arg Phe Leu Ala Ile	
65 70 75 80	
AAC TCC CAA CCC GGG CTC GAA AAT CGG GAC ACT GCC TGC GAG GAG CAA	288
Asn Ser Gln Pro Gly Leu Glu Asn Arg Asp Thr Ala Cys Glu Glu Gln	

(A) LENGTH: 834 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION:1..831

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

GTG TCA TCT GCT CCA ACC GTG TCG GTG ATA ACG ATT TCG CTG AAC GAT	48
Val Ser Ser Ala Pro Thr Val Ser Val Ile Thr Ile Ser Leu Asn Asp	
130 135 140	
CTC GAG GGA TTG AAA AGC ACC GTG GAG AGC GTT CGC GCG CAG CGC TAT	96
Leu Glu Gly Leu Lys Ser Thr Val Glu Ser Val Arg Ala Gln Arg Tyr	
145 150 155	
GGG GGG CGA ATC GAG CAC ATC GTC ATC GAC GGT GGA TCG GGC GAC GCC	144
Gly Gly Arg Ile Glu His Ile Val Ile Asp Gly Gly Ser Gly Asp Ala	
160 165 170	
GTC GTG GAG TAT CTG TCC GGC GAT CCT GGC TTT GCA TAT TGG CAA TCT	192
Val Val Glu Tyr Leu Ser Gly Asp Pro Gly Phe Ala Tyr Trp Gln Ser	
175 180 185	
CAG CCC GAC AAC GGG AGA TAT GAC GCG ATG AAT CAG GGC ATT GCC CAT	240
Gln Pro Asp Asn Gly Arg Tyr Asp Ala Met Asn Gln Gly Ile Ala His	
190 195 200 205	
TCG TCG GGC GAC CTG TTG TGG TTT ATG CAC TCC ACG GAT CGT TTC TCC	288
Ser Ser Gly Asp Leu Leu Trp Phe Met His Ser Thr Asp Arg Phe Ser	
210 215 220	
GAT CCA GAT GCA GTC GCT TCC GTG GTG GAG GCG CTC TCG GGG CAT GGA	336
Asp Pro Asp Ala Val Ala Ser Val Val Glu Ala Leu Ser Gly His Gly	
225 230 235	
CCA GTA CGT GAT TTG TGG GGT TAC GGG AAA AAC AAC CTT GTC GGA CTC	384
Pro Val Arg Asp Leu Trp Gly Tyr Gly Lys Asn Asn Leu Val Gly Leu	
240 245 250	
GAC GGC AAA CCA CTT TTC CCT CGG CCG TAC GGC TAT ATG CCG TTT AAG	432
Asp Gly Lys Pro Leu Phe Pro Arg Pro Tyr Gly Tyr Met Pro Phe Lys	
255 260 265	

ATG CGG AAA TTT CTG CTC GGC GCG ACG GTT GCG CAT CAG GCG ACA TTC	480
Met Arg Lys Phe Leu Leu Gly Ala Thr Val Ala His Gln Ala Thr Phe	
270 275 280 285	
TTC GGC GCG TCG CTG GTA GCC AAG TTG GGC GGT TAC GAT CTT GAT TTT	528
Phe Gly Ala Ser Leu Val Ala Lys Leu Gly Gly Tyr Asp Leu Asp Phe	
290 295 300	
GGA CTC GAG GCG GAC CAG CTG TTC ATC TAC CGT GCC GCA CTA ATA CGG	576
Gly Leu Glu Ala Asp Gln Leu Phe Ile Tyr Arg Ala Ala Leu Ile Arg	
305 310 315	
CCT CCC GTC ACG ATC GAC CGC GTG GTT TGC GAC TTC GAT GTC ACG GGA	624
Pro Pro Val Thr Ile Asp Arg Val Val Cys Asp Phe Asp Val Thr Gly	
320 325 330	
CCT GGT TCA ACC CAG CCC ATC CGT GAG CAC TAT CGG ACC CTG CGG CGG	672
Pro Gly Ser Thr Gln Pro Ile Arg Glu His Tyr Arg Thr Leu Arg Arg	
335 340 345	
CTC TGG GAC CTG CAT GGC GAC TAC CCG CTG GGT GGG CGC AGA GTG TCG	720
Leu Trp Asp Leu His Gly Asp Tyr Pro Leu Gly Gly Arg Arg Val Ser	
350 355 360 365	
TGG GCT TAC TTG CGT GTG AAG GAG TAC TTG ATT CGG GCC GAC CTG GCC	768
Trp Ala Tyr Leu Arg Val Lys Glu Tyr Leu Ile Arg Ala Asp Leu Ala	
370 375 380	
GCA TTC AAC GCG GTA AAG TTC TTG CGA GCG AAG TTC GCC AGA GCT TCG	816
Ala Phe Asn Ala Val Lys Phe Leu Arg Ala Lys Phe Ala Arg Ala Ser	
385 390 395	
CGG AAG CAA AAT TCA TAG	834
Arg Lys Gln Asn Ser	
400	

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

Ala Phe Asn Ala Val Lys Phe Leu Arg Ala Lys Phe Ala Arg Ala Ser
 260 265 270

Arg Lys Gln Asn Ser
 275

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1032 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1029

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

GTG AAG CGA GCG CTT ATA ACA GGG ATC ACG GGG CAG GAT GGT TCC TAC Val Lys Arg Ala Leu Ile Thr Gly Ile Thr Gly Gln Asp Gly Ser Tyr 280 285 290	48
CTC GCC GAG CTA CTA CTG AGC AAG GGA TAC GAG GTT CAC GGG CTC GTT Leu Ala Glu Leu Leu Leu Ser Lys Gly Tyr Glu Val His Gly Leu Val 295 300 305	96
CGT CGA GCT TCG ACG TTT AAC ACG TCG CGG ATC GAT CAC CTC TAC GTT Arg Arg Ala Ser Thr Phe Asn Thr Ser Arg Ile Asp His Leu Tyr Val 310 315 320 325	144
GAC CCA CAC CAA CCG GGC GCG CGC TTG TTC TTG CAC TAT GCA GAC CTC Asp Pro His Gln Pro Gly Ala Arg Leu Phe Leu His Tyr Ala Asp Leu 330 335 340	192
ACT GAC GGC ACC CGG TTG GTG ACC CTG CTC AGC AGT ATC GAC CCG GAT Thr Asp Gly Thr Arg Leu Val Thr Leu Leu Ser Ser Ile Asp Pro Asp 345 350 355	240
GAG GTC TAC AAC CTC GCA GCG CAG TCC CAT GTG CGC GTC AGC TTT GAC Glu Val Tyr Asn Leu Ala Ala Gln Ser His Val Arg Val Ser Phe Asp 360 365 370	288

GAG CCA GTG CAT ACC GGA GAC ACC ACC GGC ATG GGA TCG ATC CGA CTT Glu Pro Val His Thr Gly Asp Thr Thr Gly Met Gly Ser Ile Arg Leu 375 380 385	336
CTG GAA GCA GTC CGC CTT TCT CGG GTG GAC TGC CGG TTC TAT CAG GCT Leu Glu Ala Val Arg Leu Ser Arg Val Asp Cys Arg Phe Tyr Gln Ala 390 395 400 405	384
TCC TCG TCG GAG ATG TTC GGC GCA TCT CCG CCA CCG CAG AAC GAA TCG Ser Ser Ser Glu Met Phe Gly Ala Ser Pro Pro Pro Gln Asn Glu Ser 410 415 420	432
ACG CCG TTC TAT CCC CGT TCG CCA TAC GGC GCG GCC AAG GTC TTC TCG Thr Pro Phe Tyr Pro Arg Ser Pro Tyr Gly Ala Ala Lys Val Phe Ser 425 430 435	480
TAC TGG ACG ACT CGC AAC TAT CGA GAG GCG TAC GGA TTA TTC GCA GTG Tyr Trp Thr Thr Arg Asn Tyr Arg Glu Ala Tyr Gly Leu Phe Ala Val 440 445 450	528
AAT GGC ATC TTG TTC AAC CAT GAG TCC CCC CGG CGC GGC GAG ACT TTC Asn Gly Ile Leu Phe Asn His Glu Ser Pro Arg Arg Gly Glu Thr Phe 455 460 465	576
GTG ACC CGA AAG ATC ACG CGT GCC GTG GCG CGC ATC CGA GCT GGC GTC Val Thr Arg Lys Ile Thr Arg Ala Val Ala Arg Ile Arg Ala Gly Val 470 475 480 485	624
CAA TCG GAG GTC TAT ATG GGC AAC CTC GAT GCG ATC CGC GAC TGG GGC Gln Ser Glu Val Tyr Met Gly Asn Leu Asp Ala Ile Arg Asp Trp Gly 490 495 500	672
TAC GCG CCC GAA TAT GTC GAG GGG ATG TGG AGG ATG TTG CAA GCG CCT Tyr Ala Pro Glu Tyr Val Glu Gly Met Trp Arg Met Leu Gln Ala Pro 505 510 515	720
GAA CCT GAT GAC TAC GTC CTG GCG ACA GGG CGT GGT TAC ACC GTA CGT Glu Pro Asp Asp Tyr Val Leu Ala Thr Gly Arg Gly Tyr Thr Val Arg 520 525 530	768
GAG TTC GCT CAA GCT GCT TTT GAC CAT GTC GGG CTC GAC TGG CAA AAG Glu Phe Ala Gln Ala Ala Phe Asp His Val Gly Leu Asp Trp Gln Lys 535 540 545	816
CGC GTC AAG TTT GAC GAC CGC TAT TTG CGT CCC ACC GAG GTC GAT TCG Arg Val Lys Phe Asp Asp Arg Tyr Leu Arg Pro Thr Glu Val Asp Ser 550 555 560 565	864

CTA GTA GGA GAT GCC GAC AAG GCG GCC CAG TCA CTC GGC TGG AAA GCT	912
Leu Val Gly Asp Ala Asp Lys Ala Ala Gln Ser Leu Gly Trp Lys Ala	
570 575 580	
TCG GTT CAT ACT GGT GAA CTC GCG CGC ATC ATG GTG GAC GCG GAC ATC	960
Ser Val His Thr Gly Glu Leu Ala Arg Ile Met Val Asp Ala Asp Ile	
585 590 595	
GCC GCG TTG GAG TGC GAT GGC ACA CCA TGG ATC GAC ACG CCG ATG TTG	1008
Ala Ala Leu Glu Cys Asp Gly Thr Pro Trp Ile Asp Thr Pro Met Leu	
600 605 610	
CCT GGT TGG GGC AGA GTA AGT TGA	1032
Pro Gly Trp Gly Arg Val Ser	
615 620	

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 343 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

Val Lys Arg Ala Leu Ile Thr Gly Ile Thr Gly Gln Asp Gly Ser Tyr	
1 5 10 15	
Leu Ala Glu Leu Leu Ser Lys Gly Tyr Glu Val His Gly Leu Val	
20 25 30	
Arg Arg Ala Ser Thr Phe Asn Thr Ser Arg Ile Asp His Leu Tyr Val	
35 40 45	
Asp Pro His Gln Pro Gly Ala Arg Leu Phe Leu His Tyr Ala Asp Leu	
50 55 60	
Thr Asp Gly Thr Arg Leu Val Thr Leu Leu Ser Ser Ile Asp Pro Asp	
65 70 75 80	
Glu Val Tyr Asn Leu Ala Ala Gln Ser His Val Arg Val Ser Phe Asp	
85 90 95	
Glu Pro Val His Thr Gly Asp Thr Thr Gly Met Gly Ser Ile Arg Leu	
100 105 110	

(A) LENGTH: 1032 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

(A) NAME/KEY: CDS
 (B) LOCATION:1..1029

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

GTG AAG CGA GCG CTT ATA ACA GGG ATC ACG GGG CAG GAT GGT TCC TAC	48
Val Lys Arg Ala Leu Ile Thr Gly Ile Thr Gly Gln Asp Gly Ser Tyr	
345 350 355	
CTC GCC GAG CTA CTA CTG AGC AAG GGA TAC GAG GTT CAC GGG CTC GTT	96
Leu Ala Glu Leu Leu Leu Ser Lys Gly Tyr Glu Val His Gly Leu Val	
360 365 370 375	
CGT CGA GCT TCG ACG TTT AAC ACG TCG CGG ATC GAT CAC CTC TAC GTT	144
Arg Arg Ala Ser Thr Phe Asn Thr Ser Arg Ile Asp His Leu Tyr Val	
380 385 390	
GAC CCA CAC CAA CCG GGC GCG CGC TTG TTC TTG CAC TAT GCA GAC CTC	192
Asp Pro His Gln Pro Gly Ala Arg Leu Phe Leu His Tyr Ala Asp Leu	
395 400 405	
ACT GAC GGC ACC CGG TTG GTG ACC CTG CTC AGC AGT ATC GAC CCG GAT	240
Thr Asp Gly Thr Arg Leu Val Thr Leu Leu Ser Ser Ile Asp Pro Asp	
410 415 420	
GAG GTC TAC AAC CTC GCA GCG CAG TCC CAT GTG CGC GTC AGC TTT GAC	288
Glu Val Tyr Asn Leu Ala Ala Gln Ser His Val Arg Val Ser Phe Asp	
425 430 435	
GAG CCA GTG CAT ACC GGA GAC ACC ACC GGC ATG GGA TCG ATC CGA CTT	336
Glu Pro Val His Thr Gly Asp Thr Thr Gly Met Gly Ser Ile Arg Leu	
440 445 450 455	
CTG GAA GCA GTC CGC CTT TCT CGG GTG GAC TGC CGG TTC TAT CAG GCT	384
Leu Glu Ala Val Arg Leu Ser Arg Val Asp Cys Arg Phe Tyr Gln Ala	
460 465 470	
TCC TCG TCG GAG ATG TTC GGC GCA TCT CCG CCA CCG CAG AAC GAA TCG	432

Ala Ala Ser Glu Cys Asp Gly Thr Pro Trp Ile Asp Thr Pro Met Leu
665 670 675

CCT GGT TGG GGC GGA GTA AGT TGA
Pro Gly Trp Gly Gly Val Ser
680 685

1032

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 343 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

Val Lys Arg Ala Leu Ile Thr Gly Ile Thr Gly Gln Asp Gly Ser Tyr
1 5 10 15
Leu Ala Glu Leu Leu Leu Ser Lys Gly Tyr Glu Val His Gly Leu Val
20 25 30
Arg Arg Ala Ser Thr Phe Asn Thr Ser Arg Ile Asp His Leu Tyr Val
35 40 45
Asp Pro His Gln Pro Gly Ala Arg Leu Phe Leu His Tyr Ala Asp Leu
50 55 60
Thr Asp Gly Thr Arg Leu Val Thr Leu Leu Ser Ser Ile Asp Pro Asp
65 70 75 80
Glu Val Tyr Asn Leu Ala Ala Gln Ser His Val Arg Val Ser Phe Asp
85 90 95
Glu Pro Val His Thr Gly Asp Thr Thr Gly Met Gly Ser Ile Arg Leu
100 105 110
Leu Glu Ala Val Arg Leu Ser Arg Val Asp Cys Arg Phe Tyr Gln Ala
115 120 125
Ser Ser Ser Glu Met Phe Gly Ala Ser Pro Pro Pro Gln Asn Glu Ser
130 135 140
Thr Pro Phe Tyr Pro Arg Ser Pro Tyr Gly Ala Ala Lys Val Phe Ser
145 150 155 160

000155 000155 000155

Tyr Trp Thr Thr Arg Asn Tyr Arg Glu Ala Tyr Gly Leu Phe Ala Val
165 170 175

Asn Gly Ile Leu Phe Asn His Glu Ser Pro Arg Arg Gly Glu Thr Phe
180 185 190

Val Thr Arg Lys Ile Thr Arg Ala Val Ala Arg Ile Arg Ala Gly Val
195 200 205

Gln Ser Glu Val Tyr Met Gly Asn Leu Asp Ala Ile Arg Asp Trp Gly
210 215 220

Tyr Ala Pro Glu Tyr Val Glu Gly Met Trp Arg Met Leu Gln Ala Pro
225 230 235 240

Glu Pro Asp Asp Tyr Val Leu Ala Thr Gly Arg Gly Tyr Thr Val Arg
245 250 255

Glu Phe Ala Gln Ala Ala Phe Asp His Val Gly Leu Asp Trp Gln Lys
260 265 270

His Val Lys Phe Asp Asp Arg Tyr Leu Arg Pro Thr Glu Val Asp Ser
275 280 285

Leu Val Gly Asp Ala Asp Arg Ala Ala Gln Ser Leu Gly Trp Lys Ala
290 295 300

Ser Val His Thr Gly Glu Leu Ala Arg Ile Met Val Asp Ala Asp Ile
305 310 315 320

Ala Ala Ser Glu Cys Asp Gly Thr Pro Trp Ile Asp Thr Pro Met Leu
325 330 335

Pro Gly Trp Gly Gly Val Ser
340

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1020 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

GAC GCG TAT GCG ATC GCC AAG ATC GCC GGT ATC CTG CAA GTT CAG GCG Asp Ala Tyr Ala Ile Ala Lys Ile Ala Gly Ile Leu Gln Val Gln Ala 505 510 515	528
GTT AGG CGC CAA TAT GGG CTG GCG TGG ATC TCT GCG ATG CCG ACT AAC Val Arg Arg Gln Tyr Gly Leu Ala Trp Ile Ser Ala Met Pro Thr Asn 520 525 530 535	576
CTC TAC GGA CCC GGC GAC AAC TTC TCC CCG TCC GGG TCG CAT CTC TTG Leu Tyr Gly Pro Gly Asp Asn Phe Ser Pro Ser Gly Ser His Leu Leu 540 545 550	624
CCG GCG CTC ATC CGT CGA TAT GAG GAA GCC AAA GCT GGT GGT GCA GAA Pro Ala Leu Ile Arg Arg Tyr Glu Glu Ala Lys Ala Gly Gly Ala Glu 555 560 565	672
GAG GTG ACG AAT TGG GGG ACC GGT ACT CCG CGG CGC GAA CTT CTG CAT Glu Val Thr Asn Trp Gly Thr Gly Thr Pro Arg Arg Glu Leu Leu His 570 575 580	720
GTC GAC GAT CTG GCG AGC GCA TGC CTG TTC CTT TTG GAA CAT TTC GAT Val Asp Asp Leu Ala Ser Ala Cys Leu Phe Leu Leu Glu His Phe Asp 585 590 595	768
GGT CCG AAC CAC GTC AAC GTG GGC ACC GGC GTC GAT CAC AGC ATT AGC Gly Pro Asn His Val Asn Val Gly Thr Gly Val Asp His Ser Ile Ser 600 605 610 615	816
GAG ATC GCA GAC ATG GTC GCT ACA GCG GTG GGC TAC ATC GGC GAA ACA Glu Ile Ala Asp Met Val Ala Thr Ala Val Gly Tyr Ile Gly Glu Thr 620 625 630	864
CGT TGG GAT CCA ACT AAA CCC GAT GGA ACC CCG CGC AAA CTA TTG GAC Arg Trp Asp Pro Thr Lys Pro Asp Gly Thr Pro Arg Lys Leu Leu Asp 635 640 645	912
GTC TCC GCG CTA CGC GAG TTG GGT TGG CGC CCG CGA ATC GCA CTG AAA Val Ser Ala Leu Arg Glu Leu Gly Trp Arg Pro Arg Ile Ala Leu Lys 650 655 660	960
GAC GGC ATC GAT GCA ACG GTG TCG TGG TAC CGC ACA AAT GCC GAT GCC Asp Gly Ile Asp Ala Thr Val Ser Trp Tyr Arg Thr Asn Ala Asp Ala 665 670 675	1008
GTG AGG AGG TAA Val Arg Arg *	1020
680	

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:1..1020

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GTG CGA TGG CAC ACC ATG GAT CGA CAC GCC GAT GTT GCC TGG TTG GGG	48
Val Arg Trp His Thr Met Asp Arg His Ala Asp Val Ala Trp Leu Gly	
345 350 355	
CAG AGT AAG TTG ACG ACT ACA CCT GGG CCT CTG GAC CGC GCA ACG CCC	96
Gln Ser Lys Leu Thr Thr Thr Pro Gly Pro Leu Asp Arg Ala Thr Pro	
360 365 370 375	
GTG TAT ATC GCC GGT CAT CGG GGG CTG GTC GGC TCA GCG CTC GTA CGT	144
Val Tyr Ile Ala Gly His Arg Gly Leu Val Gly Ser Ala Leu Val Arg	
380 385 390	
AGA TTT GAG GCC GAG GGG TTC ACC AAT CTC ATT GTG CGA TCA CGC GAT	192
Arg Phe Glu Ala Glu Gly Phe Thr Asn Leu Ile Val Arg Ser Arg Asp	
395 400 405	
GAG ATT GAT CTG ACG GAC CGA GCC GCA ACG TTT GAT TTT GTG TCT GAG	240
Glu Ile Asp Leu Thr Asp Arg Ala Ala Thr Phe Asp Phe Val Ser Glu	
410 415 420	
ACA AGA CCA CAG GTG ATC ATC GAT GCG GCC GCA CGG GTC GGC GGC ATC	288
Thr Arg Pro Gln Val Ile Ile Asp Ala Ala Ala Arg Val Gly Gly Ile	
425 430 435	
ATG GCG AAT AAC ACC TAT CCC GCG GAC TTC TTG TCC GAA AAC CTC CGA	336
Met Ala Asn Asn Thr Tyr Pro Ala Asp Phe Leu Ser Glu Asn Leu Arg	
440 445 450 455	
ATC CAG ACC AAT TTG CTC GAC GCA GCT GTC GCC GTG CGT GTG CCG CGG	384
Ile Gln Thr Asn Leu Leu Asp Ala Ala Val Ala Val Arg Val Pro Arg	
460 465 470	
CTC CTT TTC CTC GGT TCG TCA TGC ATC TAC CCG AAG TAC GCT CCG CAA	432
Leu Leu Phe Leu Gly Ser Ser Cys Ile Tyr Pro Lys Tyr Ala Pro Gln	
475 480 485	
CCT ATC CAC GAG AGT GCT TTA TTG ACT GGC CCT TTG GAG CCC ACC AAC	480
Pro Ile His Glu Ser Ala Leu Leu Thr Gly Pro Leu Glu Pro Thr Asn	
490 495 500	

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 340 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

Val Arg Trp His Thr Met Asp Arg His Ala Asp Val Ala Trp Leu Gly
1 5 10 15
Gln Ser Lys Leu Thr Thr Thr Pro Gly Pro Leu Asp Arg Ala Thr Pro
20 25 30
Val Tyr Ile Ala Gly His Arg Gly Leu Val Gly Ser Ala Leu Val Arg
35 40 45
Arg Phe Glu Ala Glu Gly Phe Thr Asn Leu Ile Val Arg Ser Arg Asp
50 55 60
Glu Ile Asp Leu Thr Asp Arg Ala Ala Thr Phe Asp Phe Val Ser Glu
65 70 75 80
Thr Arg Pro Gln Val Ile Ile Asp Ala Ala Ala Arg Val Gly Gly Ile
85 90 95
Met Ala Asn Asn Thr Tyr Pro Ala Asp Phe Leu Ser Glu Asn Leu Arg
100 105 110
Ile Gln Thr Asn Leu Leu Asp Ala Ala Val Ala Val Arg Val Pro Arg
115 120 125
Leu Leu Phe Leu Gly Ser Ser Cys Ile Tyr Pro Lys Tyr Ala Pro Gln
130 135 140
Pro Ile His Glu Ser Ala Leu Leu Thr Gly Pro Leu Glu Pro Thr Asn
145 150 155 160
Asp Ala Tyr Ala Ile Ala Lys Ile Ala Gly Ile Leu Gln Val Gln Ala
165 170 175
Val Arg Arg Gln Tyr Gly Leu Ala Trp Ile Ser Ala Met Pro Thr Asn
180 185 190
Leu Tyr Gly Pro Gly Asp Asn Phe Ser Pro Ser Gly Ser His Leu Leu
195 200 205

Pro Ala Leu Ile Arg Arg Tyr Glu Glu Ala Lys Ala Gly Gly Ala Glu
210 215 220

Glu Val Thr Asn Trp Gly Thr Gly Thr Pro Arg Arg Glu Leu Leu His
225 230 235 240

Val Asp Asp Leu Ala Ser Ala Cys Leu Phe Leu Leu Glu His Phe Asp
245 250 255

Gly Pro Asn His Val Asn Val Gly Thr Gly Val Asp His Ser Ile Ser
260 265 270

Glu Ile Ala Asp Met Val Ala Thr Ala Val Gly Tyr Ile Gly Glu Thr
275 280 285

Arg Trp Asp Pro Thr Lys Pro Asp Gly Thr Pro Arg Lys Leu Leu Asp
290 295 300

Val Ser Ala Leu Arg Glu Leu Gly Trp Arg Pro Arg Ile Ala Leu Lys
305 310 315 320

Asp Gly Ile Asp Ala Thr Val Ser Trp Tyr Arg Thr Asn Ala Asp Ala
325 330 335

Val Arg Arg *
340

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1020 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1020

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GTG CGA TGG CAC ACC ATG GAT CGA CAC GCC GAT GTT GCC TGG TTG GGG
Val Arg Trp His Thr Met Asp Arg His Ala Asp Val Ala Trp Leu Gly
345 350 355

48

67

069T60"069T60"

CGG AGT AAG TTG ACG ACT ACA CCT GGG CCT CTG GAC CGC GCA ACG CCC	96
Arg Ser Lys Leu Thr Thr Thr Pro Gly Pro Leu Asp Arg Ala Thr Pro	
360 365 370	
GTG TAT ATC GCC GGT CAT CGG GGG CTG GTC GGC TCA GCG CTC GTA CGT	144
Val Tyr Ile Ala Gly His Arg Gly Leu Val Gly Ser Ala Leu Val Arg	
375 380 385	
AGA TTT GAG GCC GAG GGG TTC ACC AAT CTC ATT GTG CGA TCA CGC GAT	192
Arg Phe Glu Ala Glu Gly Phe Thr Asn Leu Ile Val Arg Ser Arg Asp	
390 395 400	
GAG ATT GAT CTG ACG GAC CGA GCC GCA ACG TTT GAT TTT GTG TCT GAG	240
Glu Ile Asp Leu Thr Asp Arg Ala Ala Thr Phe Asp Phe Val Ser Glu	
405 410 415 420	
ACA AGA CCA CAG GTG ATC ATC GAT GCG GCC GCA CGG GTC GGC GGC ATC	288
Thr Arg Pro Gln Val Ile Ile Asp Ala Ala Ala Arg Val Gly Gly Ile	
425 430 435	
ATG GCG AAT AAC ACC TAT CCC GCG GAC TTC TTG TCC GAA AAC CTC CGA	336
Met Ala Asn Asn Thr Tyr Pro Ala Asp Phe Leu Ser Glu Asn Leu Arg	
440 445 450	
ATC CAG ACC AAT TTG CTC GAC GCA GCT GTC GCC GTG CGT GTG CCG CGG	384
Ile Gln Thr Asn Leu Leu Asp Ala Ala Val Ala Val Arg Val Pro Arg	
455 460 465	
CTC CTT TTC CTC GGT TCG TCA TGC ATC TAC CCG AAG TAC GCT CCG CAA	432
Leu Leu Phe Leu Gly Ser Ser Cys Ile Tyr Pro Lys Tyr Ala Pro Gln	
470 475 480	
CCT ATC CAC GAG AGT GCT TTA TTG ACT GGC CCT TTG GAG CCC ACC AAC	480
Pro Ile His Glu Ser Ala Leu Leu Thr Gly Pro Leu Glu Pro Thr Asn	
485 490 495 500	
GAC GCG TAT GCG ATC GCC AAG ATC GCC GGT ATC CTG CAA GTT CAG GCG	528
Asp Ala Tyr Ala Ile Ala Lys Ile Ala Gly Ile Leu Gln Val Gln Ala	
505 510 515	
GTT AGG CGC CAA TAT GGG CTG GCG TGG ATC TCT GCG ATG CCG ACT AAC	576
Val Arg Arg Gln Tyr Gly Leu Ala Trp Ile Ser Ala Met Pro Thr Asn	
520 525 530	
CTC TAC GGA CCC GGC GAC AAC TTC TCC CCG TCC GGG TCG CAT CTC TTG	624
Leu Tyr Gly Pro Gly Asp Asn Phe Ser Pro Ser Gly Ser His Leu Leu	
535 540 545	

CCG GCG CTC ATC CGT CGA TAT GAG GAA GCC AAA GCT GGT GGT GCA GAA	672
Pro Ala Leu Ile Arg Arg Tyr Glu Glu Ala Lys Ala Gly Gly Ala Glu	
550 555 560	
GAG GTG ACG AAT TGG GGG ACC GGT ACT CCG CGG CGC GAA CTT CTG CAT	720
Glu Val Thr Asn Trp Gly Thr Gly Thr Pro Arg Arg Glu Leu Leu His	
565 570 575 580	
GTC GAC GAT CTG GCG AGC GCA TGC CTG TTC CTT TTG GAA CAT TTC GAT	768
Val Asp Asp Leu Ala Ser Ala Cys Leu Phe Leu Leu Glu His Phe Asp	
585 590 595	
GGT CCG AAC CAC GTC AAC GTG GGC ACC GGC GTC GAT CAC AGC ATT AGC	816
Gly Pro Asn His Val Asn Val Gly Thr Gly Val Asp His Ser Ile Ser	
600 605 610	
GAG ATC GCA GAC ATG GTC GCT ACG GCG GTG GGC TAC ATC GGC GAA ACA	864
Glu Ile Ala Asp Met Val Ala Thr Ala Val Gly Tyr Ile Gly Glu Thr	
615 620 625	
CGT TGG GAT CCA ACT AAA CCC GAT GGA ACC CCG CGC AAA CTA TTG GAC	912
Arg Trp Asp Pro Thr Lys Pro Asp Gly Thr Pro Arg Lys Leu Leu Asp	
630 635 640	
GTC TCC GCG CTA CGC GAG TTG GGT TGG CGC CCG CGA ATC GCA CTG AAA	960
Val Ser Ala Leu Arg Glu Leu Gly Trp Arg Pro Arg Ile Ala Leu Lys	
645 650 655 660	
GAC GGC ATC GAT GCA ACG GTG TCG TGG TAC CGC ACA AAT GCC GAT GCC	1008
Asp Gly Ile Asp Ala Thr Val Ser Trp Tyr Arg Thr Asn Ala Asp Ala	
665 670 675	
GTG AGG AGG TAA	1020
Val Arg Arg *	
680	

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 340 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

Val Arg Trp His Thr Met Asp Arg His Ala Asp Val Ala Trp Leu Gly
1 5 10 15

Arg Ser Lys Leu Thr Thr Thr Pro Gly Pro Leu Asp Arg Ala Thr Pro
20 25 30

Val Tyr Ile Ala Gly His Arg Gly Leu Val Gly Ser Ala Leu Val Arg
35 40 45

Arg Phe Glu Ala Glu Gly Phe Thr Asn Leu Ile Val Arg Ser Arg Asp
50 55 60

Glu Ile Asp Leu Thr Asp Arg Ala Ala Thr Phe Asp Phe Val Ser Glu
65 70 75 80

Thr Arg Pro Gln Val Ile Ile Asp Ala Ala Ala Arg Val Gly Gly Ile
85 90 95

Met Ala Asn Asn Thr Tyr Pro Ala Asp Phe Leu Ser Glu Asn Leu Arg
100 105 110

Ile Gln Thr Asn Leu Leu Asp Ala Ala Val Ala Val Arg Val Pro Arg
115 120 125

Leu Leu Phe Leu Gly Ser Ser Cys Ile Tyr Pro Lys Tyr Ala Pro Gln
130 135 140

Pro Ile His Glu Ser Ala Leu Leu Thr Gly Pro Leu Glu Pro Thr Asn
145 150 155 160

Asp Ala Tyr Ala Ile Ala Lys Ile Ala Gly Ile Leu Gln Val Gln Ala
165 170 175

Val Arg Arg Gln Tyr Gly Leu Ala Trp Ile Ser Ala Met Pro Thr Asn
180 185 190

Leu Tyr Gly Pro Gly Asp Asn Phe Ser Pro Ser Gly Ser His Leu Leu
195 200 205

Pro Ala Leu Ile Arg Arg Tyr Glu Glu Ala Lys Ala Gly Gly Ala Glu
210 215 220

Glu Val Thr Asn Trp Gly Thr Gly Thr Pro Arg Arg Glu Leu Leu His
225 230 235 240

Val Asp Asp Leu Ala Ser Ala Cys Leu Phe Leu Leu Glu His Phe Asp
245 250 255

09091538-054638

CAG TTC GGT AGC GCT TTG CGT CGT GCA GGA TTC AAG AGC CGT ATC GTT Gln Phe Gly Ser Ala Leu Arg Arg Ala Gly Phe Lys Ser Arg Ile Val 390 395 400	192
TCC TTT GAA CCT CTT TCG GGG CCA TTT GCG CAA CTA ACG CGC AAG TCG Ser Phe Glu Pro Leu Ser Gly Pro Phe Ala Gln Leu Thr Arg Lys Ser 405 410 415 420	240
GCA TCG GAT CCA CTA TGG GAG TGT CAC CAG TAT GCC CTA GGC GAC GCC Ala Ser Asp Pro Leu Trp Glu Cys His Gln Tyr Ala Leu Gly Asp Ala 425 430 435	288
GAT GAG ACG ATT ACC ATC AAT GTG GCA GGC AAT GCG GGG GCA AGT AGT Asp Glu Thr Ile Thr Ile Asn Val Ala Gly Asn Ala Gly Ala Ser Ser 440 445 450	336
TCC GTG CTG CCG ATG CTT AAA AGT CAT CAA GAT GCC TTT CCT CCC GCG Ser Val Leu Pro Met Leu Lys Ser His Gln Asp Ala Phe Pro Pro Ala 455 460 465	384
AAT TAT ATT GGC ACC GAA GAC GTT GCA ATA CAC CGC CTT GAT TCG GTT Asn Tyr Ile Gly Thr Glu Asp Val Ala Ile His Arg Leu Asp Ser Val 470 475 480	432
GCA TCA GAA TTT CTG AAC CCT ACC GAT GTT ACT TTC CTG AAG ATC GAC Ala Ser Glu Phe Leu Asn Pro Thr Asp Val Thr Phe Leu Lys Ile Asp 485 490 495 500	480
GTA CAG GGT TTC GAG AAG CAG GTT ATC ACG GGC AGT AAG TCA ACG CTT Val Gln Gly Phe Glu Lys Gln Val Ile Thr Gly Ser Lys Ser Thr Leu 505 510 515	528
AAC GAA AGC TGC GTC GGC ATG CAA CTC GAA CTT TCT TTT ATT CCG TTG Asn Glu Ser Cys Val Gly Met Gln Leu Glu Leu Ser Phe Ile Pro Leu 520 525 530	576
TAC GAA GGT GAC ATG CTG ATT CAT GAA GCG CTT GAA CTT GTC TAT TCC Tyr Glu Gly Asp Met Leu Ile His Glu Ala Leu Glu Leu Val Tyr Ser 535 540 545	624
CTA GGT TTC AGA CTG ACG GGT TTG TTG CCC GGC TTT ACG GAT CCG CGC Leu Gly Phe Arg Leu Thr Gly Leu Leu Pro Gly Phe Thr Asp Pro Arg 550 555 560	672
AAT GGT CGA ATG CTT CAA GCT GAC GGC ATT TTC TTC CGT GGG GAC GAT Asn Gly Arg Met Leu Gln Ala Asp Gly Ile Phe Phe Arg Gly Asp Asp 565 570 575 580	720

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 240 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Asp Phe Leu Arg Asn Ala Gly Leu Met Ala Arg Asn Val Ser Thr
 1 5 10 15
 Glu Met Leu Arg His Phe Glu Arg Lys Arg Leu Leu Val Asn Gln Phe
 20 25 30
 Lys Ala Tyr Gly Val Asn Val Val Ile Asp Val Gly Ala Asn Ser Gly
 35 40 45
 Gln Phe Gly Ser Ala Leu Arg Arg Ala Gly Phe Lys Ser Arg Ile Val
 50 55 60
 Ser Phe Glu Pro Leu Ser Gly Pro Phe Ala Gln Leu Thr Arg Lys Ser
 65 70 75 80
 Ala Ser Asp Pro Leu Trp Glu Cys His Gln Tyr Ala Leu Gly Asp Ala
 85 90 95
 Asp Glu Thr Ile Thr Ile Asn Val Ala Gly Asn Ala Gly Ala Ser Ser
 100 105 110
 Ser Val Leu Pro Met Leu Lys Ser His Gln Asp Ala Phe Pro Pro Ala
 115 120 125
 Asn Tyr Ile Gly Thr Glu Asp Val Ala Ile His Arg Leu Asp Ser Val
 130 135 140
 Ala Ser Glu Phe Leu Asn Pro Thr Asp Val Thr Phe Leu Lys Ile Asp
 145 150 155 160
 Val Gln Gly Phe Glu Lys Gln Val Ile Thr Gly Ser Lys Ser Thr Leu
 165 170 175
 Asn Glu Ser Cys Val Gly Met Gln Leu Glu Leu Ser Phe Ile Pro Leu
 180 185 190

Tyr Glu Gly Asp Met Leu Ile His Glu Ala Leu Glu Leu Val Tyr Ser
195 200 205

Leu Gly Phe Arg Leu Thr Gly Leu Leu Pro Gly Phe Thr Asp Pro Arg
210 215 220

Asn Gly Arg Met Leu Gln Ala Asp Gly Ile Phe Phe Arg Gly Asp Asp
225 230 235 240

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 723 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..720

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

ATG GAT TTT TTG CGC AAC GCC GGC TTG ATG GCT CGT AAC GTT AGC ACC	48
Met Asp Phe Leu Arg Asn Ala Gly Leu Met Ala Arg Asn Val Ser Thr	
245 250 255	
GAG ATG CTG CGC CAC TTC GAA CGA AAG CGC CTA TTA GTA AAC CAA TTC	96
Glu Met Leu Arg His Phe Glu Arg Lys Arg Leu Leu Val Asn Gln Phe	
260 265 270	
AAA GCA TAC GGA GTC AAC GTT GTT ATT GAT GTC GGT GCT AAC TCC GGC	144
Lys Ala Tyr Gly Val Asn Val Val Ile Asp Val Gly Ala Asn Ser Gly	
275 280 285	
CAG TTC GGT AGC GCT TTG CGT CGT GCA GGA TTC AAG AGC CGT ATC GTT	192
Gln Phe Gly Ser Ala Leu Arg Arg Ala Gly Phe Lys Ser Arg Ile Val	
290 295 300	
TCC TTT GAA CCT CTT TCG GGG CCA TTT GCG CAA CTA ACG CGC GAG TCG	240
Ser Phe Glu Pro Leu Ser Gly Pro Phe Ala Gln Leu Thr Arg Glu Ser	
305 310 315 320	

GCA TCG GAT CCA CTA TGG GAG TGT CAC CAG TAT GCC CTA GGC GAC GCC	288
Ala Ser Asp Pro Leu Trp Glu Cys His Gln Tyr Ala Leu Gly Asp Ala	
325 330 335	
GAT GAG ACG ATT ACC ATC AAT GTG GCA GGC AAT GCG GGG GCA AGT AGT	336
Asp Glu Thr Ile Thr Ile Asn Val Ala Gly Asn Ala Gly Ala Ser Ser	
340 345 350	
TCC GTG CTG CCG ATG CTT AAA AGT CAT CAA GAT GCC TTT CCT CCC GCG	384
Ser Val Leu Pro Met Leu Lys Ser His Gln Asp Ala Phe Pro Pro Ala	
355 360 365	
AAT TAT ATT GGC ACC GAA GAC GTT GCA ATA CAC CGC CTT GAT TCG GTT	432
Asn Tyr Ile Gly Thr Glu Asp Val Ala Ile His Arg Leu Asp Ser Val	
370 375 380	
GCA TCA GAA TTT CTG AAC CCT ACC GAT GTT ACT TTC CTG AAG ATC GAC	480
Ala Ser Glu Phe Leu Asn Pro Thr Asp Val Thr Phe Leu Lys Ile Asp	
385 390 395 400	
GTA CAG GGT TTC GAG AAG CAG GTT ATC GCG GGC AGT AAG TCA ACG CTT	528
Val Gln Gly Phe Glu Lys Gln Val Ile Ala Gly Ser Lys Ser Thr Leu	
405 410 415	
AAC GAA AGC TGC GTC GGC ATG CAA CTC GAA CTT TCT TTT ATT CCG TTG	576
Asn Glu Ser Cys Val Gly Met Gln Leu Glu Leu Ser Phe Ile Pro Leu	
420 425 430	
TAC GAA GGT GAC ATG CTG ATT CAT GAA GCG CTT GAA CTT GTC TAT TCC	624
Tyr Glu Gly Asp Met Leu Ile His Glu Ala Leu Glu Leu Val Tyr Ser	
435 440 445	
CTA GGT TTC AGA CTG ACG GGT TTG TTG CCC GGA TTT ACG GAT CCG CGC	672
Leu Gly Phe Arg Leu Thr Gly Leu Leu Pro Gly Phe Thr Asp Pro Arg	
450 455 460	
AAT GGT CGA ATG CTT CAA GCT GAC GGC ATT TTC TTC CGT GGG GAC GAT	720
Asn Gly Arg Met Leu Gln Ala Asp Gly Ile Phe Phe Arg Gly Asp Asp	
465 470 475 480	
TGA	723

(2) INFORMATION FOR SEQ ID NO: 20:

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

Met Asp Phe Leu Arg Asn Ala Gly Leu Met Ala Arg Asn Val Ser Thr
1 5 10 15
Glu Met Leu Arg His Phe Glu Arg Lys Arg Leu Leu Val Asn Gln Phe
20 25 30
Lys Ala Tyr Gly Val Asn Val Val Ile Asp Val Gly Ala Asn Ser Gly
35 40 45
Gln Phe Gly Ser Ala Leu Arg Arg Ala Gly Phe Lys Ser Arg Ile Val
50 55 60
Ser Phe Glu Pro Leu Ser Gly Pro Phe Ala Gln Leu Thr Arg Glu Ser
65 70 75 80
Ala Ser Asp Pro Leu Trp Glu Cys His Gln Tyr Ala Leu Gly Asp Ala
85 90 95
Asp Glu Thr Ile Thr Ile Asn Val Ala Gly Asn Ala Gly Ala Ser Ser
100 105 110
Ser Val Leu Pro Met Leu Lys Ser His Gln Asp Ala Phe Pro Pro Ala
115 120 125
Asn Tyr Ile Gly Thr Glu Asp Val Ala Ile His Arg Leu Asp Ser Val
130 135 140
Ala Ser Glu Phe Leu Asn Pro Thr Asp Val Thr Phe Leu Lys Ile Asp
145 150 155 160
Val Gln Gly Phe Glu Lys Gln Val Ile Ala Gly Ser Lys Ser Thr Leu
165 170 175
Asn Glu Ser Cys Val Gly Met Gln Leu Glu Leu Ser Phe Ile Pro Leu
180 185 190
Tyr Glu Gly Asp Met Leu Ile His Glu Ala Leu Glu Leu Val Tyr Ser
195 200 205
Leu Gly Phe Arg Leu Thr Gly Leu Leu Pro Gly Phe Thr Asp Pro Arg
210 215 220

340

345

350

CAT GCG GCA AGC CAT CTT GTC TAT GGC GAT GTT GTG ATG CGT TCG ACG 384
 His Ala Ala Ser His Leu Val Tyr Gly Asp Val Val Met Arg Ser Thr
 355 360 365

AAA AGC CGG CAT GCC GGA CCT TTC GAC CTC GAC CGC CTC CTA TTT GAG 432
 Lys Ser Arg His Ala Gly Pro Phe Asp Leu Asp Arg Leu Leu Phe Glu
 370 375 380

ACG AAT TTG TGC CAC CAA TCG ATC TTT TAC CGC CGT GAG CTT TTC GAC 480
 Thr Asn Leu Cys His Gln Ser Ile Phe Tyr Arg Arg Glu Leu Phe Asp
 385 390 395 400

GGC ATC GGC CCT TAC AAC CTG CGC TAC CGA GTC TGG GCG GAC TGG GAC 528
 Gly Ile Gly Pro Tyr Asn Leu Arg Tyr Arg Val Trp Ala Asp Trp Asp
 405 410 415

TTC AAT ATT CGC TGC TTC TCC AAC CCG GCG CTG ATT ACC CGC TAC ATG 576
 Phe Asn Ile Arg Cys Phe Ser Asn Pro Ala Leu Ile Thr Arg Tyr Met
 420 425 430

GAC GTC GTG ATT TCC GAA TAC AAC GAC ATG ACC GGC TTC AGC ATG AGG 624
 Asp Val Val Ile Ser Glu Tyr Asn Asp Met Thr Gly Phe Ser Met Arg
 435 440 445

CAG GGG ACT GAT AAA GAG TTC AGA AAA CGG CTG CCA ATG TAC TTC TGG 672
 Gln Gly Thr Asp Lys Glu Phe Arg Lys Arg Leu Pro Met Tyr Phe Trp
 450 455 460

GTT GCA GGG TGG GAG ACT TGC AGG CGC ATG CTG GCG TTT TTG AAA GAC 720
 Val Ala Gly Trp Glu Thr Cys Arg Arg Met Leu Ala Phe Leu Lys Asp
 465 470 475 480

AAG GAG AAT CGC CGT CTG GCC TTG CGT ACG CGG TTG ATA AGG GTT AAG 768
 Lys Glu Asn Arg Arg Leu Ala Leu Arg Thr Arg Leu Ile Arg Val Lys
 485 490 495

GCC GTC TCC AAA GAA CGA AGC GCA GAA CCG TAG 801
 Ala Val Ser Lys Glu Arg Ser Ala Glu Pro
 500 505

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 266 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Met Thr Ala Pro Val Phe Ser Ile Ile Ile Pro Thr Phe Asn Ala Ala
1 5 10 15
Val Thr Leu Gln Ala Cys Leu Gly Ser Ile Val Gly Gln Thr Tyr Arg
20 25 30
Glu Val Glu Val Val Leu Val Asp Gly Gly Ser Thr Asp Arg Thr Leu
35 40 45
Asp Ile Ala Asn Ser Phe Arg Pro Glu Leu Gly Ser Arg Leu Val Val
50 55 60
His Ser Gly Pro Asp Asp Gly Pro Tyr Asp Ala Met Asn Arg Gly Val
65 70 75 80
Gly Val Ala Thr Gly Glu Trp Val Leu Phe Leu Gly Ala Asp Asp Thr
85 90 95
Leu Tyr Glu Pro Thr Thr Leu Ala Gln Val Ala Ala Phe Leu Gly Asp
100 105 110
His Ala Ala Ser His Leu Val Tyr Gly Asp Val Val Met Arg Ser Thr
115 120 125
Lys Ser Arg His Ala Gly Pro Phe Asp Leu Asp Arg Leu Leu Phe Glu
130 135 140
Thr Asn Leu Cys His Gln Ser Ile Phe Tyr Arg Arg Glu Leu Phe Asp
145 150 155 160
Gly Ile Gly Pro Tyr Asn Leu Arg Tyr Arg Val Trp Ala Asp Trp Asp
165 170 175
Phe Asn Ile Arg Cys Phe Ser Asn Pro Ala Leu Ile Thr Arg Tyr Met
180 185 190
Asp Val Val Ile Ser Glu Tyr Asn Asp Met Thr Gly Phe Ser Met Arg
195 200 205
Gln Gly Thr Asp Lys Glu Phe Arg Lys Arg Leu Pro Met Tyr Phe Trp
210 215 220
Val Ala Gly Trp Glu Thr Cys Arg Arg Met Leu Ala Phe Leu Lys Asp

350																355																360																
CTC	TAC	GAA	CCA	ACC	ACG	TTG	GCC	CAG	GTA	GCC	GCT	TTT	CTC	GGC	GAC		336																															
Leu	Tyr	Glu	Pro	Thr	Thr	Leu	Ala	Gln	Val	Ala	Ala	Phe	Leu	Gly	Asp																																	
365				370				375																																								
CAT	GCG	GCA	AGC	CAT	CTT	GTC	TAT	GGC	GAT	GTT	GTG	ATG	CGT	TCG	ACG		384																															
His	Ala	Ala	Ser	His	Leu	Val	Tyr	Gly	Asp	Val	Val	Met	Arg	Ser	Thr																																	
380				385				390																																								
AAA	AGC	CGG	CAT	GCC	GGA	CCT	TTC	GAC	CTC	GAC	CGC	CTC	CTA	TTT	GAG		432																															
Lys	Ser	Arg	His	Ala	Gly	Pro	Phe	Asp	Leu	Asp	Arg	Leu	Leu	Phe	Glu																																	
395				400				405				410																																				
ACG	AAT	TTG	TGC	CAC	CAA	TCG	ATC	TTT	TAC	CGC	CGT	GAG	CTT	TTC	GAC		480																															
Thr	Asn	Leu	Cys	His	Gln	Ser	Ile	Phe	Tyr	Arg	Arg	Glu	Leu	Phe	Asp																																	
415				420				425																																								
GGC	ATC	GGC	CCT	TAC	AAC	CTG	CGC	TAC	CGA	GTC	TGG	GCG	GAC	TGG	GAC		528																															
Gly	Ile	Gly	Pro	Tyr	Asn	Leu	Arg	Tyr	Arg	Val	Trp	Ala	Asp	Trp	Asp																																	
430				435				440																																								
TTC	AAT	ATT	CGC	TGC	TTC	TCC	AAC	CCG	GCG	CTG	ATT	ACC	CGC	TAC	ATG		576																															
Phe	Asn	Ile	Arg	Cys	Phe	Ser	Asn	Pro	Ala	Leu	Ile	Thr	Arg	Tyr	Met																																	
445				450				455																																								
GAC	GTC	GTG	ATT	TCC	GAA	TAC	AAC	GAC	ATG	ACC	GGC	TTC	AGC	ATG	AGG		624																															
Asp	Val	Val	Ile	Ser	Glu	Tyr	Asn	Asp	Met	Thr	Gly	Phe	Ser	Met	Arg																																	
460				465				470																																								
CAG	GGG	ACT	GAT	AAA	GAG	TTC	AGA	AAA	CGG	CTG	CCA	ATG	TAC	TTC	TGG		672																															
Gln	Gly	Thr	Asp	Lys	Glu	Phe	Arg	Lys	Arg	Leu	Pro	Met	Tyr	Phe	Trp																																	
475				480				485				490																																				
GTT	GCA	GGG	TGG	GAG	ACT	TGC	AGG	CGC	ATG	CTG	GCG	TTT	TTG	AAA	GAC		720																															
Val	Ala	Gly	Trp	Glu	Thr	Cys	Arg	Arg	Met	Leu	Ala	Phe	Leu	Lys	Asp																																	
495				500				505																																								
AAG	GAG	AAT	CGC	CGT	CTG	GCC	TTG	CGT	ACG	CGG	TTG	ATA	AGG	GTT	AAG		768																															
Lys	Glu	Asn	Arg	Arg	Leu	Ala	Leu	Arg	Thr	Arg	Leu	Ile	Arg	Val	Lys																																	
510				515				520																																								
GCC	GTC	TCC	AAA	GAA	CGA	AGC	GCA	GAA	CCG	TAG								801																														
Ala	Val	Ser	Lys	Glu	Arg	Ser	Ala	Glu	Pro																																							
525				530																																												

81

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 266 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

Met Thr Ala Pro Val Phe Ser Ile Ile Ile Pro Thr Phe Asn Ala Ala
 1 5 10 15
 Val Thr Leu Gln Ala Cys Leu Gly Ser Ile Val Gly Gln Thr Tyr Arg
 20 25 30
 Glu Val Glu Val Val Leu Val Asp Gly Gly Ser Thr Asp Arg Thr Leu
 35 40 45
 Asp Ile Ala Asn Ser Phe Arg Pro Glu Leu Gly Ser Arg Leu Val Val
 50 55 60
 His Ser Gly Pro Asp Asp Gly Pro Tyr Asp Ala Met Asn Arg Gly Val
 65 70 75 80
 Gly Val Ala Thr Gly Glu Trp Val Leu Phe Leu Gly Ala Asp Asp Thr
 85 90 95
 Leu Tyr Glu Pro Thr Thr Leu Ala Gln Val Ala Ala Phe Leu Gly Asp
 100 105 110
 His Ala Ala Ser His Leu Val Tyr Gly Asp Val Val Met Arg Ser Thr
 115 120 125
 Lys Ser Arg His Ala Gly Pro Phe Asp Leu Asp Arg Leu Leu Phe Glu
 130 135 140
 Thr Asn Leu Cys His Gln Ser Ile Phe Tyr Arg Arg Glu Leu Phe Asp
 145 150 155 160
 Gly Ile Gly Pro Tyr Asn Leu Arg Tyr Arg Val Trp Ala Asp Trp Asp
 165 170 175
 Phe Asn Ile Arg Cys Phe Ser Asn Pro Ala Leu Ile Thr Arg Tyr Met
 180 185 190
 Asp Val Val Ile Ser Glu Tyr Asn Asp Met Thr Gly Phe Ser Met Arg
 195 200 205
 Gln Gly Thr Asp Lys Glu Phe Arg Lys Arg Leu Pro Met Tyr Phe Trp

210

215

220

Val Ala Gly Trp Glu Thr Cys Arg Arg Met Leu Ala Phe Leu Lys Asp
 225 230 235 240

Lys Glu Asn Arg Arg Leu Ala Leu Arg Thr Arg Leu Ile Arg Val Lys
 245 250 255

Ala Val Ser Lys Glu Arg Ser Ala Glu Pro
 260 265

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 867 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..864

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

GTG GCC AGC AGA AGT CCC CAC TCC GCT GCG GGT GGT TGG CTA ATT CTT	48
Val Ala Ser Arg Ser Pro His Ser Ala Ala Gly Gly Trp Leu Ile Leu	
270 275 280	
GGC GGC TCC CTT CTT GTG GTC GGC GTG GCG CAT CCG GTA GGA CTC GCC	96
Gly Gly Ser Leu Leu Val Val Gly Val Ala His Pro Val Gly Leu Ala	
285 290 295	
GGA GGT GAC GAC GAT GCT GGC GTG GTG CAG CAG CCG ATC GAG GAT GCT	144
Gly Gly Asp Asp Asp Ala Gly Val Val Gln Gln Pro Ile Glu Asp Ala	
300 305 310	
GGC GGC GGT GGT GTG CTC GGG CAG GAA TCG CCC CCA TTG TTC GAA GGG	192
Gly Gly Gly Gly Val Leu Gly Gln Glu Ser Pro Pro Leu Phe Glu Gly	
315 320 325 330	
CCA ATG CGA GGC GAT GGC CAG GGA GCG GCG CTC GTA GCC GGC AGC CAC	240
Pro Met Arg Gly Asp Gly Gln Gly Ala Ala Leu Val Ala Gly Ser His	
335 340 345	

GAG CCG GAA CAA CAG TTG AGT CCC GGT GTC GTC GAG CGG GGC GAA GCC 288
 Glu Pro Glu Gln Gln Leu Ser Pro Gly Val Val Glu Arg Gly Glu Ala
 350 355 360

GAT CTC GTC CAA GAT GAC CAG ATC CGC GCG GAG CAG GGT GTC GAT GAT 336
 Asp Leu Val Gln Asp Asp Gln Ile Arg Ala Glu Gln Gly Val Asp Asp
 365 370 375

CTT GCC GAC GGT GTT GTC GGC CAG GCC GCG GTA GAG GAC CTC GAT CAG 384
 Leu Ala Asp Gly Val Val Gly Gln Ala Ala Val Glu Asp Leu Asp Gln
 380 385 390

GTC GGC GGC GGT GAA GTA GCG GAC TTT GAA TCC GGC GTG GAC GGC AGC 432
 Val Gly Gly Gly Glu Val Ala Asp Phe Glu Ser Gly Val Asp Gly Ser
 395 400 405 410

GTG CCC GCA GCC GAT GAG CAG GTG ACT TTT GCC CGT ACC AGG TGG GCC 480
 Val Pro Ala Ala Asp Glu Gln Val Thr Phe Ala Arg Thr Arg Trp Ala
 415 420 425

AAT GAC CGC CAG GTT CTG TTG TGC CCG AAT CCA TTC CAG GCT CGA CAG 528
 Asn Asp Arg Gln Val Leu Leu Cys Pro Asn Pro Phe Gln Ala Arg Gln
 430 435 440

GTA GTC GAA CGT GGC TGC GGT GAT CGA CGA TCC GGT GAC GTC GAA CCC 576
 Val Val Glu Arg Gly Cys Gly Asp Arg Arg Ser Gly Asp Val Glu Pro
 445 450 455

GTC GAG GGT CTT GGT GAC CGG GAA GGC TGC GGC CTT GAG ACG GTT GGC 624
 Val Glu Gly Leu Gly Asp Arg Glu Gly Cys Gly Leu Glu Thr Val Gly
 460 465 470

GGT GTT GGA GGC ATC GCG GGC AGC GAT CTC GGC CTC AAC CAA CGT CCG 672
 Gly Val Gly Gly Ile Ala Gly Ser Asp Leu Gly Leu Asn Gln Arg Pro
 475 480 485 490

CAG GAT CTC CTC CGG TGT CCA GCG TTG CGT CTT GGC GAC TTG CAA CAC 720
 Gln Asp Leu Leu Arg Cys Pro Ala Leu Arg Leu Gly Asp Leu Gln His
 495 500 505

CTC GGC GGC GTT GCG GCG CAC CGT GGC CAG CTT CAA CCG CCG CAG CGC 768
 Leu Gly Gly Val Ala Ala His Arg Gly Gln Leu Gln Pro Pro Gln Arg
 510 515 520

CGC GTC AAG GTC AGC AGC CAG CGG TGC CGC CGA GGA CGG TGC CAC CGC 816
 Arg Val Lys Val Ser Ser Gln Arg Cys Arg Arg Gly Arg Cys His Arg
 525 530 535

CTT GGC AGC GGT GGT CAT GAG GCC GTC CCG TCG GTG GTG TTG ATC TTG 864
 Leu Gly Ser Gly Gly His Glu Ala Val Pro Ser Val Val Leu Ile Leu
 540 545 550

TAG 867

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 288 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Val Ala Ser Arg Ser Pro His Ser Ala Ala Gly Gly Trp Leu Ile Leu
 1 5 10 15
 Gly Gly Ser Leu Leu Val Val Gly Val Ala His Pro Val Gly Leu Ala
 20 25 30
 Gly Gly Asp Asp Asp Ala Gly Val Val Gln Gln Pro Ile Glu Asp Ala
 35 40 45
 Gly Gly Gly Gly Val Leu Gly Gln Glu Ser Pro Pro Leu Phe Glu Gly
 50 55 60
 Pro Met Arg Gly Asp Gly Gln Gly Ala Ala Leu Val Ala Gly Ser His
 65 70 75 80
 Glu Pro Glu Gln Gln Leu Ser Pro Gly Val Val Glu Arg Gly Glu Ala
 85 90 95
 Asp Leu Val Gln Asp Asp Gln Ile Arg Ala Glu Gln Gly Val Asp Asp
 100 105 110
 Leu Ala Asp Gly Val Val Gly Gln Ala Ala Val Glu Asp Leu Asp Gln
 115 120 125
 Val Gly Gly Gly Glu Val Ala Asp Phe Glu Ser Gly Val Asp Gly Ser
 130 135 140
 Val Pro Ala Ala Asp Glu Gln Val Thr Phe Ala Arg Thr Arg Trp Ala
 145 150 155 160
 Asn Asp Arg Gln Val Leu Leu Cys Pro Asn Pro Phe Gln Ala Arg Gln

165

170

175

Val Val Glu Arg Gly Cys Gly Asp Arg Arg Ser Gly Asp Val Glu Pro
 180 185 190

Val Glu Gly Leu Gly Asp Arg Glu Gly Cys Gly Leu Glu Thr Val Gly
 195 200 205

Gly Val Gly Gly Ile Ala Gly Ser Asp Leu Gly Leu Asn Gln Arg Pro
 210 215 220

Gln Asp Leu Leu Arg Cys Pro Ala Leu Arg Leu Gly Asp Leu Gln His
 225 230 235 240

Leu Gly Gly Val Ala Ala His Arg Gly Gln Leu Gln Pro Pro Gln Arg
 245 250 255

Arg Val Lys Val Ser Ser Gln Arg Cys Arg Arg Gly Arg Cys His Arg
 260 265 270

Leu Gly Ser Gly Gly His Glu Ala Val Pro Ser Val Val Leu Ile Leu
 275 280 285

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1739 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..945

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:945..1736

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

ATGGGCTGCC TCAAAGGTGG TGTCGTCGCC AATGTTGTTG TTCCAACACC GGATTATGTG

60

86

09091530 091030

CGCTCGACGA CACCGGGACT CAACTGTTGT TCCGGCTCGT GGCTGCCGGC TACGAGCGCC 1560
GCTCCCTGGC CATCGCCTCG CATTGGCCCT TCGAACAATG GGGGCGATTC CTGCCCAGAGC 1620
ACACCACCGC CGCCAGCATC CTCGATCGGC TGCTGCACCA CGCCAGCATC GTCGTACCT 1680
CCGGCGAGTC CTACCGGATG CGCCACGCCG ACCACAAGAA GGGAGCCGCC AAGAATTAG 1739

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 315 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Met Gly Cys Leu Lys Gly Gly Val Val Ala Asn Val Val Val Pro Thr
1 5 10 15
Pro Asp Tyr Val Arg Phe Ala Ser His Tyr Gly Phe Val Pro Asp Phe
20 25 30
Cys His Gly Ala Asp Pro Gln Ser Lys Gly Ile Val Glu Asn Leu Cys
35 40 45
Gly Tyr Ala Gln Asp Asp Leu Ala Val Pro Leu Leu Thr Glu Ala Ala
50 55 60
Leu Ala Gly Glu Gln Val Asp Leu Arg Ala Leu Asn Ala Gln Ala Gln
65 70 75 80
Leu Trp Cys Ala Glu Val Asn Ala Thr Val His Ser Glu Ile Cys Ala
85 90 95
Val Pro Asn Asp Arg Leu Val Asp Glu Arg Thr Val Leu Arg Glu Leu
100 105 110
Pro Ser Leu Arg Pro Thr Ile Gly Ser Gly Ser Val Arg Arg Lys Val
115 120 125
Asp Gly Leu Ser Cys Ile Arg Tyr Gly Ser Ala Arg Tyr Ser Val Pro
130 135 140
Gln Arg Leu Val Gly Ala Thr Val Ala Val Val Val Asp His Gly Ala
145 150 155 160

Leu Ile Leu Leu Glu Pro Ala Thr Gly Val Ile Val Ala Glu His Glu
165 170 175

Leu Val Ser Pro Gly Glu Val Ser Ile Leu Asp Glu His Tyr Asp Gly
180 185 190

Pro Arg Pro Ala Pro Ser Arg Gly Pro Arg Pro Lys Thr Gln Ala Glu
195 200 205

Lys Arg Phe Cys Ala Leu Gly Thr Glu Ala Gln Gln Phe Leu Val Gly
210 215 220

Ala Ala Ala Ile Gly Asn Thr Arg Leu Lys Ser Glu Leu Asp Ile Leu
225 230 235 240

Leu Gly Leu Gly Ala Ala His Gly Glu Gln Ala Leu Ile Asp Ala Leu
245 250 255

Arg Arg Ala Val Ala Phe Arg Arg Phe Arg Ala Ala Asp Val Arg Ser
260 265 270

Ile Leu Ala Ala Gly Ala Gly Thr Pro Gln Pro Arg Pro Ala Gly Asp
275 280 285

Ala Leu Val Leu Asp Leu Pro Thr Val Glu Thr Arg Ser Leu Glu Ala
290 295 300

Tyr Lys Ile Asn Thr Thr Asp Gly Thr Ala Ser
305 310 315

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 264 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Met Thr Thr Ala Ala Lys Pro Val Ala Pro Ser Ser Ala Ala Pro Leu
1 5 10 15

Ala Ala Asp Leu Asp Ala Ala Leu Arg Arg Leu Lys Leu Ala Thr Val
20 25 30

Arg Arg Asn Ala Ala Glu Val Leu Gln Val Ala Lys Thr Gln Arg Trp

45

His Lys Lys Gly Ala Ala Lys Asn
260

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 789 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

GTGACGTCTG CTCCGACCGT CTCGGTGATA ACGATCTCGT TCAACGACCT CGACGGGTTG	60
CAGCGCACGG TGAAGAGTGT GCGGGCGCAA CGCTACCGGG GACGCATCGA GCACATCGTA	120
ATCGACGGTG GCAGCGGCGA CGACGTGGTG GCATACCTGT CCGGGTGTGA ACCAGGCTTC	180
GCGTATTGGC AGTCCGAGCC CGACGGCGGG CGGTACGACG CGATGAACCA GGGCATCGCG	240
CACGCATCGG GTGATCTGTT GTGGTTCTTG CACTCCGCCG ATCGTTTTTC CGGGCCCGAC	300
GTGGTAGCCC AGGCCGTGGA GGCCTATCC GGCAAGGGAC CGGTGTCCGA ATTGTGGGGC	360
TTCGGGATGG ATCGTCTCGT CGGGCTCGAT CGGGTGCGCG GCCCGATACC TTTCAGCCTG	420
CGCAAATTCC TGGCCGGCAA GCAGGTTGTT CCGCATCAAG CATCGTTCTT CGGATCATCG	480
CTGGTGGCCA AGATCGGTGG CTACGACCTT GATTTCGGGA TCGCCGCCGA CCAGGAATTC	540
ATATTGCGGG CCGCGCTGGT ATGCGAGCCG GTCACGATTC GGTGTGTGCT GTGCGAGTTC	600
GACACCACGG GCGTCGGCTC GCACCGGGAA CCAAGCGCGG TCTTCGGTGA TCTGCGCCGC	660
ATGGGCGACC TTCATCGCCG CTACCCGTTC GGGGGAAGGC GAATATCACA TGCCTACCTA	720
CGCGGCCGGG AGTTCTACGC CTACAACAGT CGATTCTGGG AAAACGTCTT CACGCGAATG	780
TCGAAATAG	789

(2) INFORMATION FOR SEQ ID NO: 31:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 262 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Met Thr Ser Ala Pro Thr Val Ser Val Ile Thr Ile Ser Phe Asn Asp
1 5 10 15

Leu Asp Gly Leu Gln Arg Thr Val Lys Ser Val Arg Ala Gln Arg Tyr
20 25 30

Arg Gly Arg Ile Glu His Ile Val Ile Asp Gly Gly Ser Gly Asp Asp
35 40 45

Val Val Ala Tyr Leu Ser Gly Cys Glu Pro Gly Phe Ala Tyr Trp Gln
50 55 60

Ser Glu Pro Asp Gly Gly Arg Tyr Asp Ala Met Asn Gln Gly Ile Ala
65 70 75 80

His Ala Ser Gly Asp Leu Leu Trp Phe Leu His Ser Ala Asp Arg Phe
85 90 95

Ser Gly Pro Asp Val Val Ala Gln Ala Val Glu Ala Leu Ser Gly Lys
100 105 110

Gly Pro Val Ser Glu Leu Trp Gly Phe Gly Met Asp Arg Leu Val Gly
115 120 125

Leu Asp Arg Val Arg Gly Pro Ile Pro Phe Ser Leu Arg Lys Phe Leu
130 135 140

Ala Gly Lys Gln Val Val Pro His Gln Ala Ser Phe Phe Gly Ser Ser
145 150 155 160

Leu Val Ala Lys Ile Gly Gly Tyr Asp Leu Asp Phe Gly Ile Ala Ala
165 170 175

Asp Gln Glu Phe Ile Leu Arg Ala Ala Leu Val Cys Glu Pro Val Thr
180 185 190

Ile Arg Cys Val Leu Cys Glu Phe Asp Thr Thr Gly Val Gly Ser His
195 200 205

Arg Glu Pro Ser Ala Val Phe Gly Asp Leu Arg Arg Met Gly Asp Leu
210 215 220

His Arg Arg Tyr Pro Phe Gly Gly Arg Arg Ile Ser His Ala Tyr Leu
225 230 235 240

Arg Gly Arg Glu Phe Tyr Ala Tyr Asn Ser Arg Phe Trp Glu Asn Val
245 250 255

00016601660

Phe Thr Arg Met Ser Lys
260

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1023 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

GTGAAGCGAG CGCTCATCAC CGGAATCACC GGCCAGGACG GCTCGTATCT CGCCGAAC	60
CTGCTGGCCA AGGGGTATGA GGTTCACGGG CTCATCCGGC GCGCTTCGAC GTTCAACACC	120
TCGCGGATCG ATCACCTCTA CGTCGACCCG CACCAACCGG GCGCGCGGCT GTTTCTGCAC	180
TATGGTGACC TGATCGACGG AACCCGTTG GTGACCCTGC TGAGCACCAT CGAACCCGAC	240
GAGGTGTACA ACCTGGCGGC GCAGTCACAC GTGCGGGTGA GCTTCGACGA ACCCGTGCAC	300
ACCGGTGACA CCACCGGCAT GGGATCCATG CGACTGCTGG AAGCCGTTTC GCTCTCTCGG	360
GTGCACTGCC GCTTCTATCA GCGGTCCTCG TCGGAGATGT TCGGCGCCTC GCCGCCACCG	420
CAGAACGAGC TGACGCCGTT CTACCCGCGG TCACCGTATG GCGCCGCCAA GGTCTATTCG	480
TACTGGGCGA CCCGCAATTA TCGCGAAGCG TACGGATTGT TCGCCGTTAA CGGCATCTTG	540
TTCAATCACG AATCACCGCG GCGCGGTGAG ACGTTCGTGA CCCGAAAGAT CACCAGGGCC	600
GTGGCACGCA TCAAGGCCGG TATCCAGTCC GAGGTCTATA TGGGCAATCT GGATGCGGTC	660
CGCGACTGGG GGTACGCGCC CGAATACGTC GAAGGCATGT GCGGATGCT GCAGACCGAC	720
GAGCCCGACG ACTTCGTTTT GCGGACCGGG CGCGGTTTCA CCGTGCGTGA GTTCGCGCGG	780
GCCGCGTTTC AGCATGCCGG TTTGGACTGG CAGCAGTACG TGAAATTCGA CCAACGCTAT	840
CTGCGGCCCA CCGAGGTGGA TTCGCTGATC GCGGACGCGA CCAAGGCTGC CGAATTGCTG	900
GGCTGGAGGG CTTCCGGTGCA CACTGACGAG TTGGCTCGGA TCATGGTCGA CGCGGACATG	960

GCGGCGCTGG AGTGCGAAGG CAAGCCGTGG ATCGACAAGC CGATGATCGC CGGCCGGACA 1020

TGA 1023

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 340 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

Met Lys Arg Ala Leu Ile Thr Gly Ile Thr Gly Gln Asp Gly Ser Tyr
1 5 10 15
Leu Ala Glu Leu Leu Leu Ala Lys Gly Tyr Glu Val His Gly Leu Ile
20 25 30
Arg Arg Ala Ser Thr Phe Asn Thr Ser Arg Ile Asp His Leu Tyr Val
35 40 45
Asp Pro His Gln Pro Gly Ala Arg Leu Phe Leu His Tyr Gly Asp Leu
50 55 60
Ile Asp Gly Thr Arg Leu Val Thr Leu Leu Ser Thr Ile Glu Pro Asp
65 70 75 80
Glu Val Tyr Asn Leu Ala Ala Gln Ser His Val Arg Val Ser Phe Asp
85 90 95
Glu Pro Val His Thr Gly Asp Thr Thr Gly Met Gly Ser Met Arg Leu
100 105 110
Leu Glu Ala Val Arg Leu Ser Arg Val His Cys Arg Phe Tyr Gln Ala
115 120 125
Ser Ser Ser Glu Met Phe Gly Ala Ser Pro Pro Pro Gln Asn Glu Leu
130 135 140
Thr Pro Phe Tyr Pro Arg Ser Pro Tyr Gly Ala Ala Lys Val Tyr Ser
145 150 155 160
Tyr Trp Ala Thr Arg Asn Tyr Arg Glu Ala Tyr Gly Leu Phe Ala Val
165 170 175

Asn Gly Ile Leu Phe Asn His Glu Ser Pro Arg Arg Gly Glu Thr Phe
180 185 190

Val Thr Arg Lys Ile Thr Arg Ala Val Ala Arg Ile Lys Ala Gly Ile
195 200 205

Gln Ser Glu Val Tyr Met Gly Asn Leu Asp Ala Val Arg Asp Trp Gly
210 215 220

Tyr Ala Pro Glu Tyr Val Glu Gly Met Trp Arg Met Leu Gln Thr Asp
225 230 235 240

Glu Pro Asp Asp Phe Val Leu Ala Thr Gly Arg Gly Phe Thr Val Arg
245 250 255

Glu Phe Ala Arg Ala Ala Phe Glu His Ala Gly Leu Asp Trp Gln Gln
260 265 270

Tyr Val Lys Phe Asp Gln Arg Tyr Leu Arg Pro Thr Glu Val Asp Ser
275 280 285

Leu Ile Gly Asp Ala Thr Lys Ala Ala Glu Leu Leu Gly Trp Arg Ala
290 295 300

Ser Val His Thr Asp Glu Leu Ala Arg Ile Met Val Asp Ala Asp Met
305 310 315 320

Ala Ala Leu Glu Cys Glu Gly Lys Pro Trp Ile Asp Lys Pro Met Ile
325 330 335

Ala Gly Arg Thr
340

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 732 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..729

1. *Chlorophyll a* (Chl *a*)
 2. *Chlorophyll b* (Chl *b*)
 3. *Chlorophyll c* (Chl *c*)
 4. *Chlorophyll d* (Chl *d*)
 5. *Chlorophyll e* (Chl *e*)
 6. *Chlorophyll f* (Chl *f*)
 7. *Chlorophyll g* (Chl *g*)
 8. *Chlorophyll h* (Chl *h*)
 9. *Chlorophyll i* (Chl *i*)
 10. *Chlorophyll j* (Chl *j*)
 11. *Chlorophyll k* (Chl *k*)
 12. *Chlorophyll l* (Chl *l*)
 13. *Chlorophyll m* (Chl *m*)
 14. *Chlorophyll n* (Chl *n*)
 15. *Chlorophyll o* (Chl *o*)
 16. *Chlorophyll p* (Chl *p*)
 17. *Chlorophyll q* (Chl *q*)
 18. *Chlorophyll r* (Chl *r*)
 19. *Chlorophyll s* (Chl *s*)
 20. *Chlorophyll t* (Chl *t*)
 21. *Chlorophyll u* (Chl *u*)
 22. *Chlorophyll v* (Chl *v*)
 23. *Chlorophyll w* (Chl *w*)
 24. *Chlorophyll x* (Chl *x*)
 25. *Chlorophyll y* (Chl *y*)
 26. *Chlorophyll z* (Chl *z*)
 27. *Chlorophyll aa* (Chl *aa*)
 28. *Chlorophyll ab* (Chl *ab*)
 29. *Chlorophyll ac* (Chl *ac*)
 30. *Chlorophyll ad* (Chl *ad*)
 31. *Chlorophyll ae* (Chl *ae*)
 32. *Chlorophyll af* (Chl *af*)
 33. *Chlorophyll ag* (Chl *ag*)
 34. *Chlorophyll ah* (Chl *ah*)
 35. *Chlorophyll ai* (Chl *ai*)
 36. *Chlorophyll aj* (Chl *aj*)
 37. *Chlorophyll ak* (Chl *ak*)
 38. *Chlorophyll al* (Chl *al*)
 39. *Chlorophyll am* (Chl *am*)
 40. *Chlorophyll an* (Chl *an*)
 41. *Chlorophyll ao* (Chl *ao*)
 42. *Chlorophyll ap* (Chl *ap*)
 43. *Chlorophyll aq* (Chl *aq*)
 44. *Chlorophyll ar* (Chl *ar*)
 45. *Chlorophyll as* (Chl *as*)
 46. *Chlorophyll at* (Chl *at*)
 47. *Chlorophyll au* (Chl *au*)
 48. *Chlorophyll av* (Chl *av*)
 49. *Chlorophyll aw* (Chl *aw*)
 50. *Chlorophyll ax* (Chl *ax*)
 51. *Chlorophyll ay* (Chl *ay*)
 52. *Chlorophyll az* (Chl *az*)
 53. *Chlorophyll ba* (Chl *ba*)
 54. *Chlorophyll bb* (Chl *bb*)
 55. *Chlorophyll bc* (Chl *bc*)
 56. *Chlorophyll bd* (Chl *bd*)
 57. *Chlorophyll be* (Chl *be*)
 58. *Chlorophyll bf* (Chl *bf*)
 59. *Chlorophyll bg* (Chl *bg*)
 60. *Chlorophyll bh* (Chl *bh*)
 61. *Chlorophyll bi* (Chl *bi*)
 62. *Chlorophyll bj* (Chl *bj*)
 63. *Chlorophyll bk* (Chl *bk*)
 64. *Chlorophyll bl* (Chl *bl*)
 65. *Chlorophyll bm* (Chl *bm*)
 66. *Chlorophyll bn* (Chl *bn*)
 67. *Chlorophyll bo* (Chl *bo*)
 68. *Chlorophyll bp* (Chl *bp*)
 69. *Chlorophyll bq* (Chl *bq*)
 70. *Chlorophyll br* (Chl *br*)
 71. *Chlorophyll bs* (Chl *bs*)
 72. *Chlorophyll bt* (Chl *bt*)
 73. *Chlorophyll bu* (Chl *bu*)
 74. *Chlorophyll bv* (Chl *bv*)
 75. *Chlorophyll bw* (Chl *bw*)
 76. *Chlorophyll bx* (Chl *bx*)
 77. *Chlorophyll by* (Chl *by*)
 78. *Chlorophyll bz* (Chl *bz*)
 79. *Chlorophyll ca* (Chl *ca*)
 80. *Chlorophyll cb* (Chl *cb*)
 81. *Chlorophyll cc* (Chl *cc*)
 82. *Chlorophyll cd* (Chl *cd*)
 83. *Chlorophyll ce* (Chl *ce*)
 84. *Chlorophyll cf* (Chl *cf*)
 85. *Chlorophyll cg* (Chl *cg*)
 86. *Chlorophyll ch* (Chl *ch*)
 87. *Chlorophyll ci* (Chl *ci*)
 88. *Chlorophyll cj* (Chl *cj*)
 89. *Chlorophyll ck* (Chl *ck*)
 90. *Chlorophyll cl* (Chl *cl*)
 91. *Chlorophyll cm* (Chl *cm*)
 92. *Chlorophyll cn* (Chl *cn*)
 93. *Chlorophyll co* (Chl *co*)
 94. *Chlorophyll cp* (Chl *cp*)
 95. *Chlorophyll cq* (Chl *cq*)
 96. *Chlorophyll cr* (Chl *cr*)
 97. *Chlorophyll cs* (Chl *cs*)
 98. *Chlorophyll ct* (Chl *ct*)
 99. *Chlorophyll cu* (Chl *cu*)
 100. *Chlorophyll cv* (Chl *cv*)
 101. *Chlorophyll cw* (Chl *cw*)
 102. *Chlorophyll cx* (Chl *cx*)
 103. *Chlorophyll cy* (Chl *cy*)
 104. *Chlorophyll cz* (Chl *cz*)
 105. *Chlorophyll da* (Chl *da*)
 106. *Chlorophyll db* (Chl *db*)
 107. *Chlorophyll dc* (Chl *dc*)
 108. *Chlorophyll dd* (Chl *dd*)
 109. *Chlorophyll de* (Chl *de*)
 110. *Chlorophyll df* (Chl *df*)
 111. *Chlorophyll dg* (Chl *dg*)
 112. *Chlorophyll dh* (Chl *dh*)
 113. *Chlorophyll di* (Chl *di*)
 114. *Chlorophyll dj* (Chl *dj*)
 115. *Chlorophyll dk* (Chl *dk*)
 116. *Chlorophyll dl* (Chl *dl*)
 117. *Chlorophyll dm* (Chl *dm*)
 118. *Chlorophyll dn* (Chl *dn*)
 119. *Chlorophyll do* (Chl *do*)
 120. *Chlorophyll dp* (Chl *dp*)
 121. *Chlorophyll dq* (Chl *dq*)
 122. *Chlorophyll dr* (Chl *dr*)
 123. *Chlorophyll ds* (Chl *ds*)
 124. *Chlorophyll dt* (Chl *dt*)
 125. *Chlorophyll du* (Chl *du*)
 126. *Chlorophyll dv* (Chl *dv*)
 127. *Chlorophyll dw* (Chl *dw*)
 128. *Chlorophyll dx* (Chl *dx*)
 129. *Chlorophyll dy* (Chl *dy*)
 130. *Chlorophyll dz* (Chl *dz*)
 131. *Chlorophyll ea* (Chl *ea*)
 132. *Chlorophyll eb* (Chl *eb*)
 133. *Chlorophyll ec* (Chl *ec*)
 134. *Chlorophyll ed* (Chl *ed*)
 135. *Chlorophyll ee* (Chl *ee*)
 136. *Chlorophyll ef* (Chl *ef*)
 1

96

520				525				530								
CAG	CCG	TTG	TAC	GAG	GGT	GGC	ATG	CTC	ATC	CGC	GAG	GCG	CTC	GAT	CTC	624
Gln	Pro	Leu	Tyr	Glu	Gly	Gly	Met	Leu	Ile	Arg	Glu	Ala	Leu	Asp	Leu	
535				540				545								
GTG	GAT	TCG	TTG	GGC	TTT	ACG	CTC	TCG	GGA	TTG	CAA	CCC	GGT	TTC	ACC	672
Val	Asp	Ser	Leu	Gly	Phe	Thr	Leu	Ser	Gly	Leu	Gln	Pro	Gly	Phe	Thr	
550				555				560								
GAC	CCC	CGC	AAC	GGT	CGA	ATG	CTG	CAG	GCC	GAT	GGC	ATC	TTC	TTC	CGG	720
Asp	Pro	Arg	Asn	Gly	Arg	Met	Leu	Gln	Ala	Asp	Gly	Ile	Phe	Phe	Arg	
565				570				575				580				
GGC	AGC	GAT	TGA													732
Gly	Ser	Asp														

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 243 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

Met	Arg	Leu	Ala	Arg	Arg	Ala	Arg	Asn	Ile	Leu	Arg	Arg	Asn	Gly	Ile
1				5					10					15	
Glu	Val	Ser	Arg	Tyr	Phe	Ala	Glu	Leu	Asp	Trp	Glu	Arg	Asn	Phe	Leu
			20					25					30		
Arg	Gln	Leu	Gln	Ser	His	Arg	Val	Ser	Ala	Val	Leu	Asp	Val	Gly	Ala
		35					40					45			
Asn	Ser	Gly	Gln	Tyr	Ala	Arg	Gly	Leu	Arg	Gly	Ala	Gly	Phe	Ala	Gly
	50					55					60				
Arg	Ile	Val	Ser	Phe	Glu	Pro	Leu	Pro	Gly	Pro	Phe	Ala	Val	Leu	Gln
65					70					75					80
Arg	Ser	Ala	Ser	Thr	Asp	Pro	Leu	Trp	Glu	Cys	Arg	Arg	Cys	Ala	Leu
				85					90					95	
Gly	Asp	Val	Asp	Gly	Thr	Ile	Ser	Ile	Asn	Val	Ala	Gly	Asn	Glu	Gly

	100		105		110	
Ala	Ser	Ser	Ser	Val	Leu	Pro
	115				120	
Met	Leu	Lys	Arg	His	Gln	Asp
					125	Ala
Phe						
Pro	Pro	Ala	Asn	Tyr	Val	Gly
	130					135
Ala	Gln	Arg	Val	Pro	Ile	His
				140		Arg
Leu						
Asp	Ser	Val	Ala	Ala	Asp	Val
	145				150	
Leu					155	
Ile	Ala	Phe	Leu			
						160
Lys	Ile	Asp	Val	Gln	Gly	Phe
				165		
Glu	Lys	Gln	Val	Ile	Ala	Gly
						175
Asp						
Ser	Thr	Val	His	Asp	Arg	Cys
			180			
Val	Gly	Met	Gln	Leu	Glu	Leu
	185					190
Ser	Phe					
Gln	Pro	Leu	Tyr	Glu	Gly	Gly
	195					200
Met	Leu	Ile	Arg	Glu	Ala	Leu
					205	Asp
Leu						
Val	Asp	Ser	Leu	Gly	Phe	Thr
	210					215
Leu	Ser	Gly	Leu	Gln	Pro	Gly
						220
Phe	Thr					
Asp	Pro	Arg	Asn	Gly	Arg	Met
	225				230	
Leu	Gln	Ala	Asp	Gly	Ile	Phe
			235			240
Arg						
Gly	Ser	Asp				

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 732 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..729

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

GTG AAA TCG TTG AAA CTC GCT CGT TTC ATC GCG CGT AGC GCC GCC TTC

Leu Pro Leu Tyr Glu Gly Gly Met Leu Ile Pro Glu Ala Leu Asp Leu
 440 445 450

GTG TAT TCC TTG GGC TTC ACG TTG ACG GGA TTG CTG CCT TGT TTC ATT 672
 Val Tyr Ser Leu Gly Phe Thr Leu Thr Gly Leu Leu Pro Cys Phe Ile
 455 460 465

GAT GCA AAT AAT GGT CGA ATG TTG CAG GCC GAC GGC ATC TTT TTC CGC 720
 Asp Ala Asn Asn Gly Arg Met Leu Gln Ala Asp Gly Ile Phe Phe Arg
 470 475 480

GAG GAC GAT TGA 732
 Glu Asp Asp
 485

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 243 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

Val Lys Ser Leu Lys Leu Ala Arg Phe Ile Ala Arg Ser Ala Ala Phe
 1 5 10 15

Glu Val Ser Arg Arg Tyr Ser Glu Arg Asp Leu Lys His Gln Phe Val
 20 25 30

Lys Gln Leu Lys Ser Arg Arg Val Asp Val Val Phe Asp Val Gly Ala
 35 40 45

Asn Ser Gly Gln Tyr Ala Ala Gly Leu Arg Arg Ala Ala Tyr Lys Gly
 50 55 60

Arg Ile Val Ser Phe Glu Pro Leu Ser Gly Pro Phe Thr Ile Leu Glu
 65 70 75 80

Ser Lys Ala Ser Thr Asp Pro Leu Trp Asp Cys Arg Gln His Ala Leu
 85 90 95

Gly Asp Ser Asp Gly Thr Val Thr Ile Asn Ile Ala Gly Asn Ala Gly
 100 105 110

Gln Ser Ser Ser Val Leu Pro Met Leu Lys Ser His Gln Asn Ala Phe

115 120 125
 Pro Pro Ala Asn Tyr Val Gly Thr Gln Glu Ala Ser Ile His Arg Leu
 130 135 140
 Asp Ser Val Ala Pro Glu Phe Leu Gly Met Asn Gly Val Ala Phe Leu
 145 150 155 160
 Lys Val Asp Val Gln Gly Phe Glu Lys Gln Val Leu Ala Gly Gly Lys
 165 170 175
 Ser Thr Ile Asp Asp His Cys Val Gly Met Gln Leu Glu Leu Ser Phe
 180 185 190
 Leu Pro Leu Tyr Glu Gly Gly Met Leu Ile Pro Glu Ala Leu Asp Leu
 195 200 205
 Val Tyr Ser Leu Gly Phe Thr Leu Thr Gly Leu Leu Pro Cys Phe Ile
 210 215 220
 Asp Ala Asn Asn Gly Arg Met Leu Gln Ala Asp Gly Ile Phe Phe Arg
 225 230 235 240
 Glu Asp Asp

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 828 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION:1..825

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

ATG GTG CAG ACG AAA CGA TAC GCC GGC TTG ACC GCA GCT AAC ACA AAG
 Met Val Gln Thr Lys Arg Tyr Ala Gly Leu Thr Ala Ala Asn Thr Lys
 245 250 255

AAA GTC GCC ATG GCC GCA CCA ATG TTT TCG ATC ATC ATC CCC ACC TTG	96
Lys Val Ala Met Ala Ala Pro Met Phe Ser Ile Ile Ile Pro Thr Leu	
260 265 270 275	
AAC GTG GCT GCG GTA TTG CCT GCC TGC CTC GAC AGC ATC GCC CGT CAG	144
Asn Val Ala Ala Val Leu Pro Ala Cys Leu Asp Ser Ile Ala Arg Gln	
280 285 290	
ACC TGC GGT GAC TTC GAG CTG GTA CTG GTC GAC GGC GGC TCG ACG GAC	192
Thr Cys Gly Asp Phe Glu Leu Val Leu Val Asp Gly Gly Ser Thr Asp	
295 300 305	
GAA ACC CTC GAC ATC GCC AAC ATT TTC GCC CCC AAC CTC GGC GAG CGG	240
Glu Thr Leu Asp Ile Ala Asn Ile Phe Ala Pro Asn Leu Gly Glu Arg	
310 315 320	
TTG ATC ATT CAT CGC GAC ACC GAC CAG GGC GTC TAC GAC GCC ATG AAC	288
Leu Ile Ile His Arg Asp Thr Asp Gln Gly Val Tyr Asp Ala Met Asn	
325 330 335	
CGC GGC GTG GAC CTG GCC ACC GGA ACG TGG TTG CTC TTT CTG GGC GCG	336
Arg Gly Val Asp Leu Ala Thr Gly Thr Trp Leu Leu Phe Leu Gly Ala	
340 345 350 355	
GAC GAC AGC CTG TAC GAG GCT GAC ACC CTG GCG CGG GTG GCC GCC TTC	384
Asp Asp Ser Leu Tyr Glu Ala Asp Thr Leu Ala Arg Val Ala Ala Phe	
360 365 370	
ATT GGC GAA CAC GAG CCC AGC GAT CTG GTA TAT GGC GAC GTG ATC ATG	432
Ile Gly Glu His Glu Pro Ser Asp Leu Val Tyr Gly Asp Val Ile Met	
375 380 385	
CGC TCA ACC AAT TTC CGC TGG GGT GGC GCC TTC GAC CTC GAC CGT CTG	480
Arg Ser Thr Asn Phe Arg Trp Gly Gly Ala Phe Asp Leu Asp Arg Leu	
390 395 400	
TTG TTC AAG CGC AAC ATC TGC CAT CAG GCG ATC TTC TAC CGC CGC GGA	528
Leu Phe Lys Arg Asn Ile Cys His Gln Ala Ile Phe Tyr Arg Arg Gly	
405 410 415	
CTC TTC GGC ACC ATC GGT CCC TAC AAC CTC CGC TAC CGG GTC CTG GCC	576
Leu Phe Gly Thr Ile Gly Pro Tyr Asn Leu Arg Tyr Arg Val Leu Ala	
420 425 430 435	
GAC TGG GAC TTC AAT ATT CGC TGC TTT TCC AAC CCA GCG CTC GTC ACC	624
Asp Trp Asp Phe Asn Ile Arg Cys Phe Ser Asn Pro Ala Leu Val Thr	
440 445 450	

CGC TAC ATG CAC GTG GTC GTT GCA AGC TAC AAC GAA TTC GGC GGG CTC	672
Arg Tyr Met His Val Val Val Ala Ser Tyr Asn Glu Phe Gly Gly Leu	
455 460 465	
AGC AAT ACG ATC GTC GAC AAG GAG TTT TTG AAG CGG CTG CCG ATG TCC	720
Ser Asn Thr Ile Val Asp Lys Glu Phe Leu Lys Arg Leu Pro Met Ser	
470 475 480	
ACG AGA CTC GGC ATA AGG CTG GTC ATA GTT CTG GTG CGC AGG TGG CCA	768
Thr Arg Leu Gly Ile Arg Leu Val Ile Val Leu Val Arg Arg Trp Pro	
485 490 495	
AAG GTG ATC AGC AGG GCC ATG GTA ATG CGC ACC GTC ATT TCT TGG CGG	816
Lys Val Ile Ser Arg Ala Met Val Met Arg Thr Val Ile Ser Trp Arg	
500 505 510 515	
CGC CGA CGT TAG	828
Arg Arg Arg	

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 275 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

Met Val Gln Thr Lys Arg Tyr Ala Gly Leu Thr Ala Ala Asn Thr Lys	
1 5 10 15	
Lys Val Ala Met Ala Ala Pro Met Phe Ser Ile Ile Ile Pro Thr Leu	
20 25 30	
Asn Val Ala Ala Val Leu Pro Ala Cys Leu Asp Ser Ile Ala Arg Gln	
35 40 45	
Thr Cys Gly Asp Phe Glu Leu Val Leu Val Asp Gly Gly Ser Thr Asp	
50 55 60	
Glu Thr Leu Asp Ile Ala Asn Ile Phe Ala Pro Asn Leu Gly Glu Arg	
65 70 75 80	
Leu Ile Ile His Arg Asp Thr Asp Gln Gly Val Tyr Asp Ala Met Asn	
85 90 95	

Arg Gly Val Asp Leu Ala Thr Gly Thr Trp Leu Leu Phe Leu Gly Ala
100 105 110

Asp Asp Ser Leu Tyr Glu Ala Asp Thr Leu Ala Arg Val Ala Ala Phe
115 120 125

Ile Gly Glu His Glu Pro Ser Asp Leu Val Tyr Gly Asp Val Ile Met
130 135 140

Arg Ser Thr Asn Phe Arg Trp Gly Gly Ala Phe Asp Leu Asp Arg Leu
145 150 155 160

Leu Phe Lys Arg Asn Ile Cys His Gln Ala Ile Phe Tyr Arg Arg Gly
165 170 175

Leu Phe Gly Thr Ile Gly Pro Tyr Asn Leu Arg Tyr Arg Val Leu Ala
180 185 190

Asp Trp Asp Phe Asn Ile Arg Cys Phe Ser Asn Pro Ala Leu Val Thr
195 200 205

Arg Tyr Met His Val Val Val Ala Ser Tyr Asn Glu Phe Gly Gly Leu
210 215 220

Ser Asn Thr Ile Val Asp Lys Glu Phe Leu Lys Arg Leu Pro Met Ser
225 230 235 240

Thr Arg Leu Gly Ile Arg Leu Val Ile Val Leu Val Arg Arg Trp Pro
245 250 255

Lys Val Ile Ser Arg Ala Met Val Met Arg Thr Val Ile Ser Trp Arg
260 265 270

Arg Arg Arg
275

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 24 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: both
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

1. **Содержание:** 1. Введение. 2. Описание системы. 3. Требования к системе. 4. Структура системы. 5. Описание компонентов. 6. Описание процессов. 7. Описание данных. 8. Описание интерфейсов. 9. Описание тестирования. 10. Заключение.

24

105